



UNIVERSIDAD
NACIONAL
DE COLOMBIA

Modelo matemático como soporte para la planificación del transporte masivo de pasajeros aplicando una estrategia de cambio de resolución

Daniel Henao Arango

Universidad Nacional de Colombia

Facultad de Minas, Departamento de Ciencias de la Computación y de la Decisión

Medellín, Colombia

2018

Modelo matemático como soporte para la planificación del transporte masivo de pasajeros aplicando una estrategia de cambio de resolución

Daniel Henao Arango

Tesis presentada como requisito parcial para optar al título de:

Magister en Ingeniería – Ingeniería de Sistemas

Directora:

PhD. Patricia Jaramillo Álvarez

Línea de Investigación:

Investigación de Operaciones

Grupo de Investigación:

Ciencias de la Decisión

Universidad Nacional de Colombia

Facultad de Minas, Departamento de Ciencias de la Computación y de la Decisión

Medellín, Colombia

2018

A la vida, por darme la oportunidad de estudiar de nuevo. A la Universidad Nacional de Colombia, por permitirme sacar lo mejor de mí. A mi familia, por su compañía y amor en este camino. A la profesora Patricia Jaramillo, por ser una guía académica, profesional y personal. Y a Carolina, Jessica y Juan Rafael, por ser testigos y compañeros de tantas experiencias.

Resumen

En esta tesis se formula un modelo matemático de optimización para resolver de manera integrada las etapas de diseño de itinerarios y asignación de flota en un sistema de transporte aéreo de pasajeros utilizando una estrategia de cambio de resolución para disminuir el tamaño del problema resultante, en términos de la cantidad de variables de decisión y ecuaciones, así como del tiempo y de la cantidad de iteraciones requeridas para resolverlo.

Para reducir el tamaño del modelo de optimización resultante se implementa una estrategia de clusterización de datos utilizando algoritmos de Aprendizaje de Máquina e Inteligencia Artificial. Estos algoritmos permiten agrupar datos en clústers de manera no trivial, de manera que los elementos pertenecientes a cada clúster son homogéneos entre sí, y los clústers contienen elementos heterogéneos entre ellos. Así, un conjunto original de datos pasa a ser reemplazado por los centroides de los clústers encontrados.

Se desarrolla un caso de aplicación en el que, usando el modelo de optimización y la estrategia de cambio de resolución propuesta, se resuelven las dos etapas de la planeación mencionadas. Se plantea el modelo con y sin clusterización de datos y se concluye que la estrategia de clusterización, además de disminuir drásticamente el tiempo de resolución del modelo, mejora la calidad de la solución encontrada, ya que se obtiene una combinación de vuelos incluidos en el itinerario operada con un costo menor que el óptimo encontrado sin aplicar la clusterización de datos y con mejor conectividad entre ellos.

Palabras clave: optimización, clusterización, diseño de itinerarios, asignación de flota, programación.

Abstract

In this thesis, a mathematical optimization model to solve the integrated problem of itinerary design and fleet assignment in a passenger air transportation system is formulated using a change-of-scale strategy to reduce the size of the resulting problem, in terms of the number of decision variables and constraints, as well as the time and number of iterations required to solve it.

To reduce the size of the resulting model, a clustering strategy is implemented using Machine Learning and Artificial Intelligence algorithms. Such algorithms allow to group data in clusters, in a non-trivial way, so that the elements belonging to one cluster are similar among them, and the clusters contain dissimilar elements. This way, an original data set is replaced by the centroids of the clusters found.

An application case is developed to solve the mentioned integrated problem using the proposed optimization model and change-of-scale strategy. The model is solved with and without data clustering. The data clustering strategy, besides drastically reducing the resolution time of the model, improves the quality of the solution found, due to a higher flexibility to find a combination of flights included in the final itinerary with higher connectivity between them and operated with a lower cost than the optimal found without the data clustering.

Keywords: optimization, clustering, itinerary design, fleet assignment, scheduling.

Contenido

	Pág.
Resumen	VII
Abstract.....	VIII
Contenido	IX
Lista de figuras.....	XII
Lista de tablas	XIII
Introducción	1
Justificación	4
Objetivos.....	7
Objetivo general	7
Objetivos específicos	7
1. Marco teórico sobre modelos de optimización para el diseño de sistemas de transporte público urbanos y aéreos.....	9
1.1 Generalidades sobre la planeación de la operación de sistemas de transporte público	9
1.2 Modelos de optimización para la planeación de la operación del transporte aéreo comercial de pasajeros	12
1.2.1 Contexto general	12
1.2.2 Integración entre etapas en la planeación de la operación del transporte aéreo	14
1.2.3 Modelo integrado de definición de itinerarios y de asignación de flota	17
1.2.4 Descripción verbal del modelo integrado de optimización para el Diseño de Itinerarios y la Asignación de Tipo de Flota	22
1.2.5 Modelo matemático explícito para usar como base para proponer el cambio de resolución.....	24
1.2.6 Preguntas abiertas en investigación en modelos de optimización para la planeación de la operación del transporte aéreo	26
2. Marco teórico sobre herramientas de Inteligencia Artificial para agrupar datos	29
2.1 Introducción	29
2.2 Características básicas del aprendizaje supervisado y no supervisado	31
2.3 Aplicaciones generales de clusterización para reducir la cantidad de datos.....	34

- X Modelo matemático como soporte para la planificación del transporte masivo de pasajeros aplicando una estrategia de cambio de resolución

3. Modelo propuesto de solución integrada de los problemas de diseño de itinerarios y de asignación de flota utilizando una estrategia de cambio de resolución.....	39
3.1 Introducción	39
3.2 Suposiciones	40
3.2.1 Entradas del modelo	40
3.2.2 Salidas del modelo	41
3.3 Formulación algebraica del modelo propuesto	41
3.3.1 Conjuntos <i>iniciales</i>	41
3.3.2 Parámetros para crear copias de cada vuelo	42
3.3.3 Notación de los vuelos copia	43
3.3.4 Otros conjuntos	43
3.3.5 Parámetros.....	44
3.3.6 Variables de decisión	45
3.3.7 Función objetivo	46
3.3.8 Restricciones.....	46
3.4 Planteamiento de un caso básico de estudio	49
3.4.1 Datos del caso básico de estudio	49
3.4.2 Programación del caso básico de estudio	56
3.4.3 Resultados del caso básico de estudio.....	57
3.4.4 Tamaño del modelo.....	58
3.5 Algoritmo propuesto de clusterización de datos	58
3.5.1 Clusterización del conjunto F de tipos de avión.....	58
3.5.2 Algoritmo de clusterización	59
3.6 Adaptación del modelo propuesto a la clusterización del conjunto F.....	61
3.6.1 Cálculo de parámetros que dependen del tipo de avión	61
3.6.2 Otros conjuntos y parámetros que no se alteraron debido a la clusterización .	61
3.7 Estrategia de reconversión de resultados para asignar a cada vuelo elegido en el itinerario un tipo específico de avión y planteamiento de un modelo reducido de asignación	62
3.7.1 Formulación de la estrategia de reconversión de resultados.....	62
3.7.2 Formulación algebraica del modelo reducido de optimización	64
 4. Caso de aplicación del modelo de optimización y de la estrategia de clusterización propuesta.....	67
4.1 Introducción.....	67
4.2 Descripción general del caso de aplicación	67
4.3 Suposiciones, adaptaciones y simplificaciones del itinerario para la formulación de caso de aplicación	68
4.4 Formulación algebraica del caso de aplicación sin involucrar clusterización de datos	70
4.5 Programación del modelo de optimización sin clusterización.....	75
4.6 Resultados del caso de aplicación sin clusterización	75
4.6.1 Resultados de las variables de decisión.....	76
4.6.2 Tamaño del modelo sin clusterización.....	77
4.7 Formulación algebraica del caso de aplicación involucrando clusterización de datos	77
4.7.1 Procedimiento de clusterización del conjunto F de tipos de avión	78
4.7.2 Cálculo de parámetros que dependen del tipo de avión	79

4.8 Programación del modelo de optimización con clusterización de los tipos de aviones	80
4.9 Resultados del caso de aplicación con clusterización	80
4.9.1 Valores de las variables de decisión.....	80
4.9.2 Tamaño del modelo involucrando clusterización	81
4.10 Implementación del cambio de escala para asignar a cada vuelo elegido un tipo de avión	82
4.11 Comparación general de resultados.....	83
5. Conclusiones y trabajos futuros.....	89
6. Bibliografía	97

Lista de figuras

	Pág.
Figura 0-1: Interacción entre la funcionalidad y la resolución de un modelo de optimización	6
Figura 1-1: Esquema general de las etapas de planeación de la operación en una aerolínea	12
Figura 1-2: Estructuras de red en una aerolínea	13
Figura 1-3: Tipos de parámetros en el modelo integrado de diseño de itinerarios y asignación de tipo de flota.....	20
Figura 1-4: Esquematización de un itinerario entre el par MDE-CTG	21
Figura 1-5: Esquematización de una red en el tiempo y en el espacio	23
Figura 2-1: Ventajas y desventajas de la clasificación k-NN.....	32
Figura 2-2: Ventajas y desventajas de la clasificación k-NN.....	34
Figura 3-1: Esquematización de la notación de los vuelos copia.....	43
Figura 3-2: Aeropuertos considerados en el caso básico de estudio.....	49
Figura 3-3: Tipos de aviones disponibles en el caso básico de estudio.....	50
Figura 3-4: Ejemplo de la notación de vuelos copia en el caso básico de estudio	51
Figura 3-5: Síntesis de la secuencia de pasos para clusterizar conjuntos de datos	60
Figura 4-1: Recorte del itinerario de la aerolínea Avianca del otoño de 2010	68
Figura 4-2: Destinos nacionales del caso de aplicación	70
Figura 4-3: Destinos internacionales del caso de aplicación	70
Figura 4-4: Rutas consideradas en el caso de aplicación.....	71
Figura 4-5: Rutas internacionales a Centroamérica, Norteamérica y El Caribe consideradas en el caso de aplicación.....	71
Figura 4-6: Rutas internacionales a Suramérica y Europa consideradas en el caso de aplicación.....	71

Lista de tablas

	Pág.
Tabla 2-1: Características del aprendizaje supervisado y no supervisado	31
Tabla 3-1: Vuelos originales del caso básico de estudio.....	50
Tabla 3-2 Conjunto ampliado de vuelos para el caso básico de estudio.....	52
Tabla 3-3: Distancias entre las ciudades consideradas en el caso básico de estudio.....	54
Tabla 3-4 Costos operativos calculados para el caso básico de estudio	54
Tabla 3-5: Rentabilidades calculadas para el caso básico de estudio	55
Tabla 3-6 Hora de salida de cada vuelo en el caso básico de estudio.....	55
Tabla 3-7 Hora de llegada de cada vuelo en el caso básico de estudio.....	55
Tabla 3-8 Aeropuerto de salida de cada vuelo en el caso básico de estudio	56
Tabla 3-9 Aeropuerto de llegada de cada vuelo en el caso básico de estudio	56
Tabla 3-10: Resultados del caso básico de estudio.....	57
Tabla 3-11 Estadísticas del modelo usando los datos del caso básico de estudio.....	58
Tabla 3-12: Conjuntos y parámetros esenciales del modelo reducido de optimización para desclusterizar las respuestas	62
Tabla 3-13: Otros conjuntos y parámetros del modelo reducido de optimización para desclusterizar las respuestas	63
Tabla 4-1: Datos técnicos de los aviones que se usarán para la clusterización	72
Tabla 4-2: Número de aviones disponibles por tipo en el caso de aplicación	74
Tabla 4-3: Resultados del caso de aplicación sin clusterizar los datos	76
Tabla 4-4: Discriminación de costos de la solución óptima del caso de aplicación sin clusterizar los datos.....	76
Tabla 4-5: Estadísticas del modelo usando los datos del caso básico de aplicación	77
Tabla 4-6: Centroides de los clústers de tipos de avión resultantes.....	79
Tabla 4-7: Número de aviones disponibles por clúster	79
Tabla 4-8: Resultados del caso de aplicación al clusterizar los datos.....	81
Tabla 4-9: Comparación entre las estadísticas del modelo usando los datos del caso básico de aplicación con y sin clusterización de datos	82
Tabla 4-10: Estadísticas de los modelos reducidos planteados para desclusterizar la solución agregada	83
Tabla 4-11: Resultados del caso de aplicación al desclusterizar los tipos de avión según los resultados de los modelos reducidos	83

XIV Modelo matemático como soporte para la planificación del transporte masivo
de pasajeros aplicando una estrategia de cambio de resolución

Tabla 4-12: Costos operativos, pérdida de rentabilidad y penalización por conexiones infactibles en los modelos con y sin clusterización	84
Tabla 4-13: Comparación de los tamaños de los modelos con y sin clusterización	85
Tabla 4-14: Similitud entre las horas de salida y los tipos de avión asignados respecto al itinerario original de Avianca	85
Tabla 4-15: Similitud entre las horas de salida y los tipos de avión asignados en el modelo con clusterización vs el modelo sin clusterización	86
Tabla 4-16: Número de conexiones exitosas establecidas por los modelos con y sin clusterización	86
Tabla 4-17: Número de vuelos asignados a cada tipo de avión en los modelos con y sin clusterización	87

Introducción

El transporte es una actividad inherente a la cotidianidad del hombre. El desplazamiento de objetos, carga o de las mismas personas, desde una ubicación de origen hasta una de destino ha permitido a lo largo de la historia de la humanidad intercambiar bienes y conectar sociedades y culturas.

Dentro de las múltiples clasificaciones que pueden hacerse sobre los tipos de transporte para facilitar su estudio, comprensión y modelización, el transporte público de pasajeros es una categoría que ha recibido especial atención en la medida en que las ciudades han ido creciendo, su población ha ido aumentando, y a que cada vez la comunicación entre empresas, países y culturas exige mayor eficiencia para mover a una persona o a una carga de un punto A hasta un punto B. En este contexto de exigencia de rapidez y de un mejor aprovechamiento de los recursos, el diseño de sistemas de transporte público de pasajeros ha cobrado importancia como un área de estudio crucial en la sociedad moderna (Guihaire & Hao, 2008), tanto por su relevancia práctica como por la diversidad de temas de investigación que se pueden tratar al respecto.

Cuando se habla de sistemas de transporte público, normalmente se piensa en los sistemas urbanos de metro, buses, trenes y tranvías. Sin embargo, el transporte aéreo de pasajeros también es una categoría de transporte público, en la medida en que muchos usuarios comparten el medio de transporte, pagan una tarifa y se someten al horario dispuesto por el operador. En este trabajo se abordará conceptualmente el transporte público de pasajeros en su generalidad, y se dará una aplicación en el contexto del transporte aéreo.

El diseño de sistemas de transporte público colectivo es un área de estudio en la que confluye el uso de diferentes herramientas de modelización, ya que su complejidad exige una aproximación holística (Poorman, 2005). Una adecuada planeación de un sistema de transporte trae consecuencias positivas tanto para los usuarios como para los operadores.

En este sentido, los usuarios que utilizan un sistema de transporte planeado y administrado gozan de una mejor calidad de vida (Lee & Sener, 2016), mientras que los operadores pueden hacer un uso más eficiente de sus recursos (infraestructura, vehículos y personal) y así garantizan una mayor rentabilidad de la operación.

En general, un problema matemático de optimización busca asignar los valores de unas variables de decisión que cumplan un conjunto dado de restricciones, de forma tal que se maximice o minimice una función objetivo. Cada una de las etapas en las que puede dividirse el proceso de planeación de la operación de sistemas de transporte público (diseño de líneas, determinación de frecuencias, diseño de horarios, programación de vehículos y programación de conductores) puede verse como un problema donde hay que tomar decisiones que satisfagan ciertas restricciones y que optimicen uno o varios objetivos. En el Capítulo 1 se verá que tanto la planeación estratégica como la planeación de la operación de sistemas de transporte se han modelizado matemáticamente en diferentes etapas sucesivas (debido a su complejidad inherente y al alto costo computacional que implicaría trabajar con un único modelo completo que abarque todas las etapas), para para una de las cuales se han formulado diversas propuestas de modelos de optimización.

A pesar de todos los avances que se han presentado desde la segunda mitad del siglo XX en todas las técnicas de optimización matemática, aplicarlas a la planificación del transporte público colectivo ha sido una tarea no trivial y que enfrenta a los planeadores a un entorno de complejidad, generado principalmente por la escala de los problemas que se desean resolver, el número de variables de decisión de diferente naturaleza que se hace necesario definir, la incertidumbre en los parámetros de los modelos, los intereses conflictivos entre los agentes involucrados en la operación, la amplia variedad de combinaciones factibles que pueden darse, entre otros. Para ilustrar el difícil avance de las técnicas de optimización en el transporte, Ceder (2007) relata que hasta finales de los años 90, muchas agencias de tránsito en Estados Unidos planeaban la operación de sus sistemas de transporte con base en un manual impreso en 1947 y reimpresso en 1982 sin mayores modificaciones. Jaramillo y Lotero (2010) por su parte explican que al modelizar la planificación de un sistema de transporte público colectivo se encuentra un problema complejo tanto analítica como computacionalmente. La complejidad analítica está dada por la gran cantidad de abstracciones y simplificaciones que hay que hacer para poder formular

modelos que representen la operación del sistema en el marco de un modelo de optimización. Un ejemplo es que los operadores y los usuarios pueden tener objetivos diferentes (y con frecuencia en conflicto) respecto a las características de operación del sistema. El operador puede desear minimizar el número de frecuencias diarias de una ruta mientras que el usuario desea minimizar el tiempo total de viaje. El analista se enfrenta a la pregunta de cómo balancear esos dos objetivos, teniendo en cuenta además que los modelos clásicos de optimización permiten involucrar una única función objetivo. Por otro lado, la complejidad computacional está dada por las características combinatoriales que tienen los modelos típicamente utilizados para modelar la planificación de los sistemas de transporte, y que generan un espectro de soluciones de gran tamaño que sería imposible solucionar por técnicas de enumeración exhaustiva de posibilidades.

Ante este panorama de complejidad, el camino más usual que han tomado los investigadores en el tema ha sido el de abordar las etapas de planificación de la operación de los sistemas de transporte desde una perspectiva secuencial, puesto que, según Guihaire & Hao (2008), abordar la modelización desde una perspectiva global absoluta resulta imposible dada su excepcional complejidad. De esta manera, cada etapa se analiza de manera independiente a las demás, generando una cascada de modelos en los que las salidas de los modelos de la parte superior son entradas para los modelos de la parte inferior. Esta partición conceptual, que tiene por objetivo transformar un gran modelo integrado en varios modelos separados, tiene la ventaja de que arroja soluciones óptimas para cada etapa en menos tiempo que lo que tardaría en hacerlo un modelo integrado. Sin embargo, el hecho de generar óptimos locales para cada una de las etapas, al final termina produciendo una solución que puede desviarse de un óptimo global, o incluso no ser factible. Un modelo integrado, que en teoría arrojaría la solución óptima global para todas las etapas de planeación, es virtualmente imposible de resolver debido a su complejidad combinatorial. Surge entonces el interrogante de explorar las estrategias matemáticas de formulación que permitan capturar las mejores características de ambos extremos de la modelización: la solución rápida de los modelos secuenciales y la búsqueda de la optimalidad de los modelos integrados.

La propuesta que se plantea en este trabajo tiene que ver con la formulación de un modelo de optimización que integre dos etapas de la planeación de la operación de sistemas de transporte público de pasajeros (con miras a robustecer una solución óptima global y no a

agregar soluciones óptimas locales independientes de cada etapa) aplicando una estrategia de cambio de resolución para disminuir su complejidad combinatorial, en términos de la cantidad de variables de decisión involucradas. Esta estrategia de cambio de escala se inspira en técnicas de clusterización que dividen de manera no trivial un conjunto de datos en subconjuntos homogéneos en su interior y heterogéneos entre ellos, de manera que, al representar un subconjunto de datos por su centroide, se disminuya la cantidad total de parámetros. El propósito de adaptar esta idea para un modelo de optimización es reducir el tamaño del espacio de búsqueda, de manera que se puedan modelizar de manera eficiente las decisiones asociadas a las dos etapas de la planeación que se desean integrar, primero en una escala de decisión *gruesa* (con un nivel de detalle agregado) y luego en una escala *fin*a, donde ya las decisiones se expresan en términos de las preguntas que originalmente se desean resolver. De esta manera se facilita el proceso de solución del modelo, al evitar optimizar un sistema tan grande y con tantos procesos involucradas con un nivel de detalle.

Justificación

Puede decirse que en la literatura hay cierto consenso sobre la utilidad de los modelos secuenciales para obtener soluciones óptimas para cada una de las etapas en las que puede dividirse la planeación de la operación de sistemas de transporte público: diseño de líneas, determinación de frecuencias, diseño de horarios, programación de vehículos y programación de conductores. Sin embargo, también se ha reconocido que agregar un conjunto de soluciones óptimas locales puede no producir una solución óptima global (Rodrigues, De Souza, & Moura, 2006; Xie, Kliwer, & Suhl, 2012). El enfoque secuencial del modelado se ha mantenido para desarrollar, en cada una de las etapas, diversidad de aproximaciones en cuanto a la definición de los modelos, así como variaciones en los enfoques de la función objetivo. Por otro lado, y a pesar de reconocer que los enfoques completamente integrados de todas las etapas de la planeación serían virtualmente imposibles de resolver por un software (debido a su complejidad combinatorial), el enfoque integrado de modelado se ha mantenido en la literatura, y ha permitido el desarrollo de diferentes temas de investigación, como el de la optimización multiobjetivo de la asignación de horarios y de vehículos (Ibarra-Rojas, Giesen, & Rios-Solis, 2014), o el desarrollo de una herramienta de asignación de conductores que con un solo modelo resuelve los dos

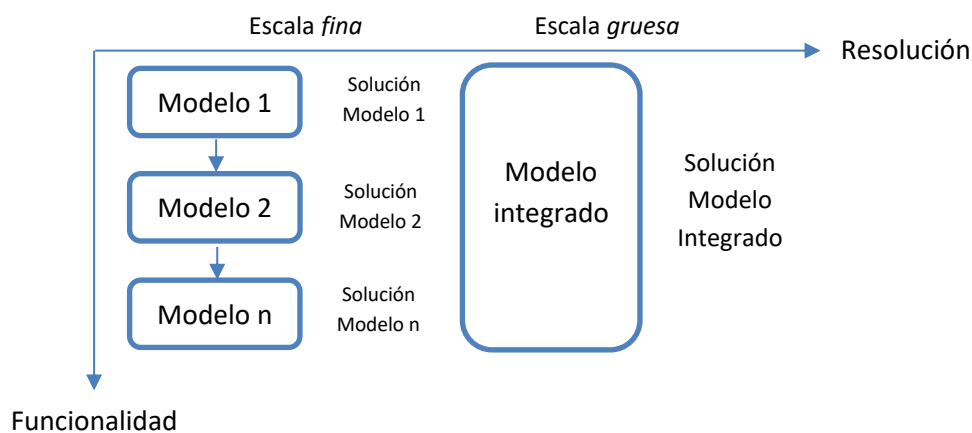
problemas típicos (crew scheduling y crew rostering) en los que ese problema se ha dividido (Xie et al., 2012).

Por otro lado, aún de forma incipiente y con cierto énfasis en las etapas de la fase estratégica de planeación, han empezado a surgir algunos modelos que desarrollan cambios de escala o de resolución como una estrategia para capturar la complejidad del sistema que se está modelando. Como ejemplos se mencionan aplicaciones de programación binivel con diferentes enfoques: optimizar el nivel de desempeño de la red de transporte y la satisfacción de uso (Migdalas, 1995); obtener una combinación óptima de carriles exclusivos en una red de transporte (Mesbah, Sarvi, Ouveysi, & Currie, 2011) u optimizar el precio de la tarifa de uso de un sistema urbano de trenes, con la premisa de satisfacer un estándar de calidad deseado (Xueyu & Jiaqi, 2013). En el trabajo de Migdalas (1995) el cambio de resolución se hace al desagregar los agentes del problema en dos niveles: un nivel de "gobierno" o planeación, que busca optimizar el desempeño de la red, y un nivel de usuarios, que toma decisiones en cuanto a las rutas y modos de transporte. La función objetivo global del problema depende de las interacciones entre dichas decisiones. En este mismo sentido, Mesbah et al. (2011) consideran dos niveles de optimización: uno superior (administrador del sistema) y uno inferior que considera partición modal, asignación de tráfico y asignación de tránsito. Igualmente, Xueyu & Jiaqi (2013) plantean un modelo de optimización a dos niveles para optimizar la tarifa del sistema de trenes: uno para la compañía de transporte y otro para los usuarios. Como se puede ver, en estos trabajos, el cambio de resolución está dado en términos de agentes: esto permite modelizar individualmente las decisiones de cada uno, pero no reduce el tamaño global del problema. Finalmente, se resaltan algunos trabajos en los que el cambio de escala se hace en el horizonte de planeación de una etapa de modelado, como los desarrollados por Marín y Jaramillo (2008), para planear la expansión de la capacidad de sistemas ferroviarios, y por (Bešinović et al., 2016; Schlechte et al., 2011), quienes generan tablas de tiempo robustas y estables para el sector ferroviario. En este último trabajo se optimiza primero el corto plazo, con una perspectiva casi de simulación, y luego, con base en las decisiones del corto plazo, se cubre el largo plazo.

Desde la perspectiva de la formulación de modelos de optimización, desarrollar una estrategia matemática que involucre un cambio en la resolución del modelo como una forma de sortear las dificultades que conlleva trabajar exclusivamente con modelos

secuenciales o integrados, puede verse como un aporte al cuerpo de conocimiento que ya existe sobre la formulación y solución de cada una de las etapas del proceso de planeación de la operación de sistemas de transporte público. Este postulado va en línea con lo expresado por Guihaire & Hao (2008), cuando afirman que ante el panorama de complejidad en el modelado de sistemas de transporte público, que impide obtener una solución global con un único modelo, dos posibles estrategias a desarrollar son el planteamiento de métodos innovadores de solución de los modelos y subdivisiones pertinentes del gran problema a considerar. La estrategia de cambio de escala puede verse como una manera de subdividir el problema para reducir su complejidad combinatorial. En la Figura 0-1 puede observarse la interacción entre la resolución de un modelo (en términos de la escala que se use) y su funcionalidad (en términos de las soluciones obtenidas). Usando una escala fina, en la que ni los datos se agrupan ni los modelos se integran, se obtiene una cadena secuencial de soluciones. Al usar una escala gruesa, en la que los modelos se integran, se obtiene una solución global para todos los modelos que se deseen resolver. La integración implica, como se ha mencionado, un aumento combinatorial del problema a optimizar, y es por esto que para subsanar el costo que esta implica, y para garantizar una solución en un tiempo razonable, se propone hacer una agrupación de los datos en clústers utilizando algoritmos de Inteligencia Artificial que, por medio de técnicas estadísticas, permitan reducir la cantidad de variables de decisión que se deban implementar en la formulación del modelo integrado, sin caer en clasificaciones triviales de los datos.

Figura 0-1: Interacción entre la funcionalidad y la resolución de un modelo de optimización



Elaboración propia

Objetivos

Objetivo general

Desarrollar un modelo de optimización para la planificación del transporte público colectivo de pasajeros que integre dos etapas de la planeación de la operación de sistemas de transporte público y que utilice una estrategia de cambio de escala o resolución en su formulación, en términos de la cantidad de variables de decisión involucradas.

Objetivos específicos

1. Formular un modelo matemático de optimización que integre las etapas de diseño de itinerarios y de asignación de tipo de flota de la planeación de la operación de sistemas de transporte público basado en los modelos más recientes en la literatura y que involucre un cambio de resolución entre elementos de escala gruesa y escala fina (en términos de cantidad de variables de decisión involucradas).
2. Comparar las variables de escala gruesa y escala fina del modelo resultante a través de una estrategia algorítmica que permita hacer el tránsito entre una escala y otra.
3. Aplicar el modelo desarrollado en el objetivo 1 para un caso de estudio reportado en la literatura, comparando las variables del caso de estudio con las propuestas en el modelo de cambio de escala, y contrastando los resultados obtenidos en cada modelo.

1.Marco teórico sobre modelos de optimización para el diseño de sistemas de transporte público urbanos y aéreos

1.1 Generalidades sobre la planeación de la operación de sistemas de transporte público

El interés de este trabajo se centra en la formulación de modelos de optimización que permitan diseñar una operación eficiente tanto para los usuarios como para los operadores, y por ello la aproximación elegida para construir el marco teórico involucra modelos matemáticos para la planeación estratégica y operativa de la operación de dichos sistemas.

El diseño y la planificación de la operación de sistemas de transporte público es un proceso caracterizado por una alta complejidad. Para sortearla desde una perspectiva matemática, el proceso de modelado de estos sistemas se ha dividido en dos grandes etapas o fases: la fase de planeación estratégica y la fase de planeación operativa (Jaramillo & Lotero, 2010). Cada una de estas fases está compuesta por sub-fases relacionadas, de manera que las salidas o resultados de una sub-fase i son parte de las entradas para la sub-fase $i+1$. El propósito de esta división es facilitar, simplificar y, en cierto sentido, linealizar el modelado, de manera que para el analista resulte más conveniente el trabajo, pero vale la pena mencionar que este es un recurso artificioso ya que al final el sistema de transporte operará como un todo.

En general, en la fase de planeación estratégica se intenta caracterizar el sistema de transporte público desde una perspectiva de largo plazo: dadas unas características sociales, económicas y geográficas de una región, se estima la cantidad de viajes que se dan desde cada posible origen hasta cada posible destino de la región bajo estudio, y se asignan estos viajes a cada uno de los diferentes modos de transporte que el sistema

contemple. Los orígenes y destinos dependen de los usos del suelo, de la ubicación de centros productivos, educativos, culturales, deportivos, y en general de los patrones de movilidad de la región. En este sentido, el diseño de un conjunto de modos de transporte que sea utilizado por los usuarios para moverse desde un origen hasta un destino implica un análisis que contemple posibles cambios en los usos del suelo y en los patrones de movilidad de la población, y por ello esta fase es de corte descriptivo, y suele utilizar herramientas de pronóstico y simulación para su modelado. Este trabajo asumirá como dados los resultados de la fase estratégica y se concentrará en los modelos de la fase operativa.

Para cada uno de los modos de transporte considerados en la fase estratégica puede desarrollarse la fase operativa. Como su nombre lo indica, el alcance de las decisiones tomadas en esta etapa está limitado al corto plazo, ya que la planeación de la operación puede estar sujeta a eventos externos o a estacionalidades que ameriten ajustes periódicos. En la fase operativa, el interés no es tanto de caracterización sino de optimización del sistema (considerando uno o múltiples objetivos), por lo que este trabajo se centrará en los modelos formulados en cada una de las etapas. Una breve descripción dada por Guihaire & Hao (2008) y Jaramillo & Lotero (2010) de cada una de las etapas que conforman la fase operativa es la siguiente:

- **Diseño de líneas:** es el problema de decidir el trazado y los recorridos del sistema. En esta etapa se obtiene un conjunto de rutas y las paradas de cada una. Las restricciones de esta etapa tienen que ver con la cobertura geográfica deseada de la red, la demanda que se desea satisfacer y la longitud de los trazados. El objetivo que se persigue, entre otros, puede ser obtener una red con un trazado que siga alguna forma geométrica.
- **Determinación de frecuencias:** asignación de la cantidad de salidas programadas que cada ruta tendrá desde su estación de origen. Las restricciones de la formulación de esta etapa tienen que ver con evitar sobresaturar la red en ciertas horas y evitar tiempos *largos* entre salidas consecutivas de una ruta. El objetivo depende de la perspectiva con que se analice: el operador desea minimizar la cantidad de frecuencias asignadas a cada ruta; el usuario, por su

parte, busca que esa cantidad sea la mayor posible, debido a la comodidad que gana teniendo más horarios disponibles.

- **Diseño de horarios:** generación de una tabla de tiempos que determine la hora de salida de cada frecuencia de cada ruta desde su estación de origen, la hora de pasada por cada estación intermedia, y la hora de llegada a la estación final. El objetivo en esta etapa puede ser minimizar el tiempo total de viaje de los usuarios o minimizar los tiempos de espera de los usuarios para hacer transferencia de una línea a otra.
- **Programación de vehículos:** asignación de vehículos específicos a cada una de las salidas de cada una de las rutas. Las restricciones de esta etapa buscan que todos los viajes programados sean cubiertos. El objetivo puede escribirse como la maximización de la utilización de la flota disponible.
- **Programación de conductores:** asignación de conductores a cada vehículo, respetando su jornada laboral. Las restricciones de esta última etapa buscan que cada vehículo tenga un conductor asignado. El objetivo se escribe como la minimización de los costos de administración de personal.

Esta desagregación de la fase operativa, como ya se ha mencionado, es un recurso artificioso que se utiliza para facilitar el modelado y el análisis del comportamiento del sistema. Ceder (2007) por ejemplo, excluye la determinación de frecuencias como una sub etapa independiente sino que la incluye dentro del desarrollo de la tabla de horarios. Por su parte, en (Bazargan, 2010) se consideran las etapas para la optimización operativa en el contexto aeronáutico y únicamente se consideran las fases de diseño de itinerarios, programación de vehículos (dividida en dos etapas: asignación de un tipo de flota a cada vuelo y luego la asignación de un avión específico a cada vuelo) y la programación de tripulaciones. La característica común de todas estas clasificaciones es que generalmente se consideran los modelos de manera secuencial, en el sentido de que las salidas de un modelo i hacen parte de las entradas del modelo $i+1$. Como ya se mencionó, esta división, si bien contribuye a facilitar el análisis y a disminuir la complejidad combinatorial de la formulación, puede conducir a soluciones que no sean óptimas globalmente. En este aspecto, Guihaire & Hao (2008) afirman que incluso al considerar estas etapas de manera independiente, los problemas que resultan son de carácter NP-hard.

1.2 Modelos de optimización para la planeación de la operación del transporte aéreo comercial de pasajeros

1.2.1 Contexto general

Como se vio en la sección anterior, la planeación de la operación de los sistemas de transporte público de pasajeros se ha dividido generalmente en etapas secuenciales: diseño de líneas, determinación de frecuencias, diseño de horarios, programación de vehículos y programación de conductores (Ceder, 2007). La industria del transporte aéreo de pasajeros ha sido típicamente tratada de manera independiente en la academia debido a que guarda importantes diferencias respecto a los sistemas urbanos de transporte público: una de ellas radica en que el transporte aéreo de pasajeros es una industria de carácter eminentemente privada, cuyo servicio consiste en ofrecer sillas para viajar de un origen a un destino. Por ser una industria privada, con la particularidad de que las sillas se consideran productos perecederos (una silla no vendida se convierte en una pérdida de ganancias a la hora de salida del vuelo (Sherali, Bish, & Zhu, 2006)), es que los esfuerzos para planear la operación se han concentrado en la rentabilidad y en la captura de la mayor demanda posible. En esta industria, de manera similar al transporte urbano, la planeación de la operación de una aerolínea, conocida en inglés como *Airline Schedule Planning*, generalmente también se ha dividido en etapas sucesivas (Lohatepanont, 2002), que se muestran en la Figura 1-1.

Figura 1-1: Esquema general de las etapas de planeación de la operación en una aerolínea



Elaboración propia con base en (Lohatepanont, 2002)

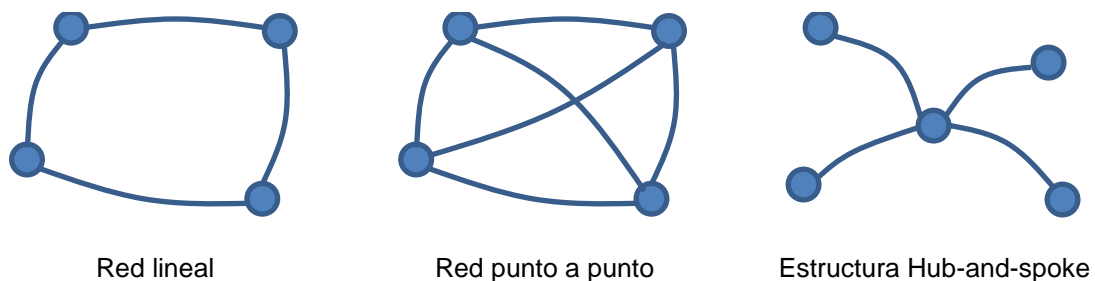
Cuando una aerolínea va a diseñar su itinerario, identifica los posibles mercados que atenderá. Un mercado se define como un par origen-destino, que tiene una demanda potencial que se desea capturar. El mercado entre dos puntos A y B no necesariamente

es el mismo que entre los puntos B y A, dado que la demanda entre las ciudades puede ser distinta en cada sentido. Así, por ejemplo, el mercado Bogotá-Medellín es diferente al mercado Medellín-Bogotá.

La primera etapa de la planeación de la operación de una aerolínea se conoce como **Diseño de Itinerarios**, y consiste en decidir a qué destinos volar y en qué horarios, de manera que se maximice la rentabilidad. En esta etapa se suelen decidir las frecuencias de vuelo de cada ruta, y el producto final es un itinerario (un conjunto de vuelos que conectan los mercados que la aerolínea busca atender). Con base en el itinerario, se plantea la segunda etapa de decisión: **Asignación de Flota**. Consiste en asignarle a cada vuelo un tipo de avión. Esta decisión depende de la demanda potencial de cada vuelo y de los costos de operación de cada tipo de avión, por lo que la asignación que se hace busca reducir los costos operativos. Una vez se sabe qué tipo de avión cubrirá cada ruta, sigue la tercera etapa de decisión, que se conoce como **Ruteo de Aviones**, y consiste en asignar, para cada uno de los aviones que conforman la flota de la aerolínea, la secuencia de vuelos que hará cada día. Cuando todos los vuelos han sido asignados a todos los aviones, la última etapa es la **Asignación de Tripulaciones** a cada vuelo. Cuidándose de cumplir todas las regulaciones de horas de trabajo permitidas para cada tripulante, es necesario asignar las tripulaciones a cada vuelo (Lohatepanont, 2002).

Cuando se elige el itinerario que la aerolínea desarrollará, se define la **estructura de red** de la aerolínea. La definición de dicha estructura tiene profundos impactos en la forma como se diseña la operación y se administran los desajustes normales en el día a día (retrasos y cancelaciones). Las tres principales estructuras se presentan en la Figura 1-2.

Figura 1-2: Estructuras de red en una aerolínea



Elaboración propia con base en (Lohatepanont, 2002)

Como se aprecia en la Figura 1-2, la red lineal es aquella con forma de circuito en la que un avión parte de una ciudad, visita secuencialmente las ciudades *vecinas* y finalmente regresa a la ciudad de origen. Por su parte, la estructura de red punto a punto, ampliamente usada por aerolíneas de bajo costo, es aquella en la que (casi todos) los destinos están conectados directamente entre sí, por lo que un pasajero rara vez debe hacer conexión para viajar de una ciudad a otra, ya que existen vuelos directos entre ellas. Esta estructura de red permite ahorrar los costos de administrar conexiones: puertas de abordaje, manejo de equipajes, posibles pérdidas de conexiones, entre otros. Finalmente, la estructura de *Hub-and-spoke* es la utilizada por la gran mayoría de aerolíneas tradicionales, en la que hay uno o varios grandes centros de conexiones (*hubs*), a donde salen y llegan los vuelos desde los distintos destinos de la aerolínea (Lohatepanont, 2002).

1.2.2 Integración entre etapas en la planeación de la operación del transporte aéreo

Para cada una de las etapas de la planeación en una aerolínea es posible formular modelos de optimización, dado que en cada etapa se deben tomar decisiones que cumplan un conjunto de restricciones y que optimicen alguna medida de desempeño (por ejemplo, la rentabilidad o el costo). Sin embargo, y de manera análoga a lo que sucede en la planeación del transporte público urbano, trabajar con modelos secuenciales puede llevar a una solución final subóptima o incluso no factible. Es por ello que luego de plantear modelos de optimización independientes para cada etapa, la investigación ha avanzado hacia integraciones entre etapas. Eso significa que si se integran, por ejemplo, dos etapas, con un solo conjunto de variables de decisión y de restricciones, se obtiene una solución factible (óptima, por supuesto) para dos de las etapas de decisión que se desean resolver, reduciendo los reprocesos y ajustes que harían falta en caso de trabajar de manera secuencial. Además, Gu, Johnson, Nemhauser, & Yinhua (1994) probaron que para más de tres tipos de aviones, el problema de asignación de flota (sin incluir el diseño de itinerarios) es un problema NP-hard, por lo que, si se desean integrar estas dos etapas, hace falta algún tipo de heurísticas que simplifiquen en algún sentido el tamaño del problema. El cambio de escala que se propone en esta tesis, agrupando datos que sean similares entre sí, es una heurística que contribuye a disminuir la complejidad (por lo menos desde la perspectiva combinatorial).

En este punto, donde se pone de manifiesto la importancia práctica de investigar sobre integraciones entre etapas en la planeación de la operación de una aerolínea, vale la pena resaltar que la investigación académica sobre modelos de optimización aplicados al transporte aéreo no ha tenido un avance tan veloz como el de la investigación en transporte público urbano. En este sentido, Akartunal, Boland, Evans, Wallace, & Waterer (2013) señalan que aunque la optimización ha hecho un aporte significativo en la planeación aeronáutica, el área sufre de una carencia de datos estandarizados y problemas benchmark, lo que en la práctica dificulta la comparación de algoritmos, heurísticas y modelos. La explicación que los autores dan a la dificultad de generar problemas benchmark puede resumirse en dos factores: por un lado, la dificultad de representar las elecciones de los viajeros y en recoger datos adecuados que permitan valorar la ganancia de un itinerario y, por otro lado, que la estructura de la red de rutas de una aerolínea depende de factores como las posiciones geográficas de los aeropuertos servidos por la red, la estructura de red de la competencia y de la demanda de pasajeros (Akartunal et al., 2013). Su investigación (que excede el alcance de esta tesis) consiste en proponer un marco y una metodología para generar datos benchmark de demanda de pasajeros, así como la estimación de parámetros de los modelos, que es de las principales limitaciones en la investigación. Un último aspecto que debe tenerse en cuenta para explicar la poca disponibilidad de datos es que, por tratarse de una industria conformada mayoritariamente por empresas privadas, no es común que haya información pública disponible sobre los procesos de planeación de itinerarios y rutas, así como datos de costos operativos. Las consecuencias prácticas de estas dificultades para soportar la investigación se han manifestado en que en la industria, en muchos casos, la planeación de la operación se continúa haciendo de manera secuencial, e incluso manual (Barnhart & Cohn, 2004).

A pesar de las dificultades para generar problemas benchmark, hay variedad de propuestas de integraciones entre etapas (trabajadas en gran medida en colaboración con aerolíneas). Algunas formulaciones consideran como variable binaria un vuelo con su hora de salida ya definido (debido a razones de mercadeo o captura de demanda). Bajo este enfoque, la estructura matemática de los problemas se asemeja a la del Problema de la Mochila: dado un conjunto de elementos factibles, se debe elegir la mejor combinación. Como se sabe, aunque esta idea suena intuitivamente rápida de ejecutar, el Problema de la Mochila está clasificado, en su complejidad de solución, como un problema NP-hard.

Para evitar un problema tipo mochila, algunas pocas formulaciones consideran que la hora a la que sale el vuelo también debe ser una variable de decisión. El hecho de que la gran mayoría de propuestas de modelos de diseño de itinerarios consideren cada vuelo con su hora de salida ya fija, marca una de las diferencias fundamentales (y posiblemente explica la diferencia en la cantidad de artículos de investigación) entre la planeación de la operación del transporte urbano y aéreo. En el primero, las etapas de determinación de frecuencias y diseño de horarios buscan calcular, para cada ruta y de manera óptima, la cantidad de viajes en el horizonte de planeación, y la hora a la que cada uno debe salir. La elección del número de viajes (frecuencia) busca optimizar el tamaño de flota requerido para satisfacer la demanda del sistema, mientras que la determinación de la hora de salida busca maximizar las posibilidades de sincronización de los usuarios en las distintas estaciones de trasbordo del sistema, de manera que el tiempo de espera para pasar de una línea a otra sea el más corto posible (Jaramillo & Lotero, 2010). Cuando en la planeación de la operación de una aerolínea, en la etapa de diseño de itinerarios, las entradas del modelo incluyen como datos fijos la cantidad de vuelos entre cada par de ciudades y la hora de salida de cada uno de ellos, se está negando desde el principio la posibilidad de optimizar las posibilidades de sincronización de los pasajeros en los aeropuertos hub, así como de capturar la mayor cantidad posible de pasajeros en las horas de mayor demanda. En este sentido, Cadarso & de Celis (2017) expresan que incrementar el número de frecuencias de un par origen-destino reduce el tiempo de espera entre vuelos y hace más atractivo volar para los pasajeros que deben elegir entre viajar por tierra o en avión. Igualmente, al aumentar la cantidad de frecuencias, se aumenta la porción de mercado que la aerolínea puede capturar en comparación con sus competidores. Respecto a la definición de la hora de salida de cada vuelo, Belobaba, Odoni, & Barnhart (2009) sostienen que una correcta elección de esta influye en aspectos tan diversos como críticos como la utilización de la flota (horas de vuelo al día), tiempo entre vuelos consecutivos, conveniencia del itinerario para los pasajeros, tiempo de abastecimiento de combustible y limpieza de los aviones, y el tiempo de conexión en aeropuertos hub. Al tomar como fijas la cantidad de vuelos y la hora de salida de cada uno, la optimización de las otras etapas de decisión puede resultar subóptima en comparación con una planeación que desde el principio considerara como variable de decisión la cantidad de vuelos entre cada par origen-destino, así como la hora de salida y llegada de cada vuelo.

A pesar de las dificultades para la formulación de modelos matemáticos para el diseño de itinerarios, las principales conclusiones a las que llegan las distintas propuestas (con base en los modelos donde se consideran fijos los horarios de los vuelos) son que los resultados de la optimización son mejores que los de una planeación manual (en términos de rentabilidad o costos, por ejemplo) y que la complejidad de los problemas se ve reflejada en el tiempo de la solución (16-19 horas) (Lohatepanont, 2002). Como recomendación, los distintos autores indican que, para el transporte aéreo, es de suma importancia considerar modelos robustos, ya que, desde la perspectiva de la operación diaria (susceptible a requerir reajustes respecto al itinerario original) es preferible una solución que funcione adecuadamente no únicamente en el caso ideal, sino que sea flexible ante situaciones inesperadas no controladas durante la etapa de planeación (clima, retrasos, ...) (Cadarso & Marín, 2012). Sin embargo, en los artículos presentados, no hacen una comparación de los resultados de integrar las etapas versus las optimizaciones secuenciales de cada etapa. Esta comparación sería una medida del desempeño de la integración entre etapas como una alternativa para superar la complejidad de la optimización secuencial.

1.2.3 Modelo integrado de definición de itinerarios y de asignación de flota

En este trabajo se abordará la integración entre las etapas de Diseño de Itinerarios y Asignación de Flota. Estas dos etapas son las más comúnmente integradas en la literatura académica y existe por tanto un amplio conjunto de investigadores que han trabajado en estos dos problemas de manera conjunta. De acuerdo con Sherali et al., (2006), la planeación de mantenimiento y la administración de ingresos (*Revenue Management*) dependen fuertemente del itinerario y del tipo de avión asignado a cada vuelo, por lo que una solución óptima de estas dos etapas tiene innegable relevancia en la industria. En este mismo sentido se expresan Dong, Chuhang, & Lau (2016), cuando afirman que entre todas las etapas, el diseño de itinerarios y la asignación de flota son las más importantes para determinar la rentabilidad y el nivel de servicio.

La suposición inicial del modelo integrado de diseño de itinerarios y asignación de flota es que se diseña un itinerario para un día, y este itinerario se repite los demás días de la semana. Asumir que la demanda es la misma todos los días de la semana es una simplificación útil pero no realista en el diseño de itinerarios (Bélanger, Desaulniers,

Soumis, Desrosiers, & Lavigne, 2006). Una aproximación sugerida (en caso de implementar estos modelos en un contexto práctico) es resolver el problema para un solo día de la semana, como un punto de partida para resolver el problema de una semana completa. Esto debido a que pueden apreciarse diferencias significativas en la demanda en días laborales en comparación con los días de fin de semana (Bélanger et al., 2006). Sin embargo, la flexibilidad de permitir asignar a un vuelo distintos tipos de avión en distintos días de la semana incrementa la complejidad computacional del problema de optimización. En consecuencia, la gran mayoría de estudios considera que un mismo tipo de avión se asigna a los vuelos de toda la semana (Sherali et al., 2006).

En general, también se asume que el diseño de itinerarios es incremental; es decir, que los itinerarios no se construyen desde cero, sino que se parte de un itinerario base (un itinerario de la temporada actual o pasada) y con base en vuelos añadidos o eliminados se llega a un nuevo itinerario. Algunas de las razones que dan quienes defienden esta práctica, según (Lohatepanont, 2002) son:

- Un itinerario completamente nuevo requiere muchos datos que pueden no estar disponibles.
- Crear un itinerario completamente nuevo es operacionalmente impráctico y computacionalmente difícil.
- Cambiar estructuras de red requiere inversiones en infraestructura aeroportuaria (puertas de abordaje, puestos de check-in, counters, salas de espera, ...)
- Las aerolíneas prefieren consistencia entre una temporada y otra, especialmente en los mercados de negocios, en los que la confiabilidad y consistencia son altamente valorados.

De acuerdo con Lohatepanont (2002), las ventajas de modificar un itinerario base son:

- Los datos históricos de reservas y pronósticos de tráfico están disponibles y pueden ser de utilidad.
- Las inversiones fijas en infraestructura de aeropuertos pueden usarse eficientemente.
- La consistencia puede mantenerse introduciendo un número limitado de cambios en el itinerario base.

En este punto, y a pesar de las razones y ventajas mencionadas, debe tenerse en cuenta que, si bien modificar un itinerario base puede ayudar a reflejar tendencias del mercado, sesga la modelización del itinerario al iniciar con una solución prefijada que brinda muy

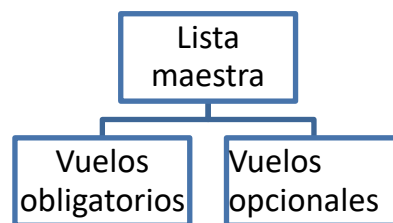
poca flexibilidad de cambios; es decir, partir de un itinerario subóptimo puede no resultar de ayuda significativa en la optimización de la rentabilidad en comparación con un itinerario construido desde cero, especialmente cuando la economía o la demografía del país que se sirve sufren cambios significativos y que afectan la demanda de pasajeros (Brons, Pels, Nijkamp, & Rietveld, 2002; Pai, 2010). En las diferentes propuestas revisadas para integrar etapas en la planeación del transporte aéreo, se encontró que normalmente primero se diseña el itinerario y posteriormente se evalúa la demanda capturada por este, utilizando un software, por ejemplo *Sabre*. Este programa, con base en el itinerario propuesto y en los itinerarios conocidos de la competencia, calcula la demanda de pasajeros que se puede capturar. (Lohatepanont (2002) argumenta que obtener detalles del funcionamiento interno de ese modelo evaluador de itinerarios ha sido difícil dada la propiedad privada sobre él). Una visión sistemática debería proponer un procedimiento invertido y cíclico: primero hacer una estimación de la demanda, luego proponer un itinerario que se acomode a ella, y finalmente evaluar la captura de demanda de ese itinerario. En la revisión de literatura realizada sobre los modelos de optimización que integran las etapas de diseño de itinerarios y asignación de flota se encontró que un modelo en el que diversos autores se han basado en los años recientes para proponer refinaciones es el de Lohatepanont (2002). Por ejemplo, Yan, Tang, & Fu (2008) consideran la demanda estocástica para elegir cuáles vuelos incluir en el itinerario, y así incluir en el resultado final las variaciones diarias en la demanda de cada vuelo. Por su parte, Pita, Barnhart, & Antunes (2012) consideran el modelo integrado de itinerarios y flota e involucran en él la congestión en los aeropuertos como un criterio para decidir cuáles vuelos elegir en el itinerario final, buscando satisfacer condiciones de slots (cantidad de operaciones permitidas en un horizonte de tiempo) en los aeropuertos que sufren mayor congestión. Por otro lado, el modelo más reciente encontrado corresponde al propuesto por Abdelghany, Abdelghany, & Azadian (2017). En su propuesta se sigue mencionando que las etapas de planeación de la operación en la industria aérea se dan de manera secuencial y que son altamente interdependientes. Enfatizan en que muchos de los parámetros que se necesitan para los modelos son inciertos y se estiman con datos históricos. En su revisión de literatura, Abdelghany et al. (2017) insisten en que la literatura se ha concentrado en proponer soluciones incrementales y no itinerarios completamente nuevos. Sin embargo, reconocen que este enfoque no necesariamente provee una solución óptima global, ya que modificar ligeramente un itinerario existente lo único que logra es una mejora marginal en el itinerario

base. Un punto adicional que destacan es que la hora de despegue de los vuelos, que se toma mayoritariamente como fija, no se optimiza en los modelos de optimización en los que se basan debido a la intratabilidad del problema resultante. En su propuesta, los autores generan una ventana de tiempo para cada vuelo (discretizando el tiempo de manera arbitraria alrededor de una hora de despegue *semilla*) y eligiendo la *mejor* hora de despegue en esa ventana.

Entradas del modelo

Los datos que se ingresan al modelo como parámetros son un conjunto de vuelos, llamados lista maestra. Los vuelos son clasificados en obligatorios y opcionales, según la obligatoriedad de ser incluidos en el itinerario final resultante.

Figura 1-3: Tipos de parámetros en el modelo integrado de diseño de itinerarios y asignación de tipo de flota.



Elaboración propia con base en (Lohatepanont, 2002)

Los vuelos obligatorios, como su nombre lo indica, son los vuelos que deben ser incluidos en el itinerario final. Por su parte, los vuelos opcionales son los vuelos que pueden o no ser incluidos en el itinerario final, bien sea porque son vuelos que actualmente existen y que pueden ser eliminados, o potenciales nuevos vuelos para ser añadidos al itinerario. Naturalmente, también se ingresan como parámetros los tipos de avión con los que cuenta la aerolínea, la cantidad de aviones de cada tipo y una estimación del costo de operación de cada vuelo en cada tipo de avión.

Salidas del modelo

Una vez se ha corrido el modelo de optimización, el resultado que se obtiene es una lista recomendada de vuelos a incluir en el itinerario final, con un tipo de avión asociado a cada vuelo. Dicha lista recomendada de vuelos es la que permite maximizar la rentabilidad de la operación y reducir los costos de operación (Lohatepanont, 2002). Sin embargo, un

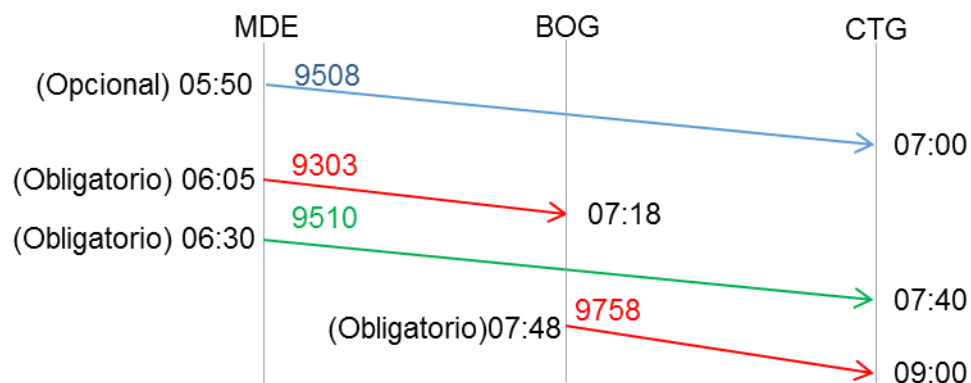
modelo inicial que optimice la rentabilidad y los costos, puede modificarse para maximizar la homogeneidad en la asignación de un tipo de avión a un vuelo (Bélanger et al., 2006). Dicha propiedad normalmente no se tiene en cuenta en el diseño de itinerarios y es posible que no maximice la rentabilidad, pero sí ayuda a incrementar la eficiencia operacional en la planeación de la operación de servicios en tierra. Como se ve, las distintas funciones objetivo que aparecen en este modelo dan cuenta de la complejidad a la que se enfrentan los planeadores de itinerarios, que incluye los múltiples objetivos que pueden perseguirse al diseñar un itinerario, las aerolíneas competidoras que buscan también capturar la mayor porción del mercado posible, la demanda estocástica (y la forma en que los pasajeros eligen en qué aerolínea volar y a qué hora) y condiciones inciertas de operación (por ejemplo el clima) (Cadarsó & de Celis, 2017). La experiencia en la industria señala que en un escenario ideal, un itinerario debe contribuir a la rentabilidad de la aerolínea, siendo viable desde la perspectiva operacional (es decir, que dé lugar a los ajustes del día a día dadas las cancelaciones y retrasos que normalmente afectan el plan original).

Definiciones

Vuelo: despegue de un origen y aterrizaje en un destino.

Itinerario: secuencia de uno o más vuelos entre un par origen-destino (Sherali et al., 2006).

Figura 1-4: Esquematzación de un itinerario entre el par MDE-CTG



Elaboración propia con base en (Sherali et al., 2006)

Como se aprecia en la Figura 1-4, un itinerario puede incluir tanto vuelos directos (9508 y 9510) como vuelos en conexión. En el ejemplo, hay pasajeros que hacen conexión entre el vuelo 9303 y 9758.

1.2.4 Descripción verbal del modelo integrado de optimización para el Diseño de Itinerarios y la Asignación de Tipo de Flota

El modelo de optimización propuesto por Sherali et al. (2006) es una simplificación del propuesto por Lohatepanont (2002). Este último plantea un modelo que incluye muchos aspectos de oferta y demanda y que afectan la rentabilidad de cada vuelo. Sherali et al. (2006) simplifican estos aspectos y por tanto presentan un modelo más sencillo de replicar.

Variables de decisión: Las variables de decisión del modelo planteado buscan responder la pregunta *¿cuáles vuelos volar y en qué tipo de avión?* Nótese que en el planteamiento de las variables ya se nota la integración entre el Diseño de Itinerarios y la Asignación de Flota, puesto que con una sola variable de decisión se resuelven las dos preguntas. La forma matemática de plantear una variable que responda esas preguntas es a través de variables binarias. La formulación de modelos de planeación usando variables binarias es común tanto en los modelos de transporte urbano como de transporte aéreo. Esta es una de las causas por las que estos modelos son especialmente difíciles de solucionar, dado que un modelo binario generalmente tiene una complejidad computacional mayor que la de un modelo lineal.

Función objetivo: Maximizar (Ganancias – Costos). Las ganancias se maximizan al elegir volar los vuelos que más contribuyen a garantizar la rentabilidad. Los costos se minimizan al asignar los tipos de avión a los vuelos de forma que cada vuelo quede con el tipo de avión *más apropiado*.

Restricciones: Sherali et al. (2006) plantean los siguientes grupos de restricciones, que también aparecen cuando se plantean modelos independientes para cada etapa:

- Restricciones de balance: Estas restricciones aseguran que cada tipo de avión sea asignado de forma lógica a cada vuelo.
- Cubrimiento de vuelos: Con estas restricciones se garantiza que:
 - Los vuelos obligatorios son incluidos dentro del itinerario.
 - Los vuelos opcionales pueden o no ser incluidos dentro del itinerario.
 - Cada vuelo elegido es asignado a un único tipo de avión.
- Disponibilidad de aviones: Estas restricciones aseguran que no se usen más aviones de los que cuenta la aerolínea.

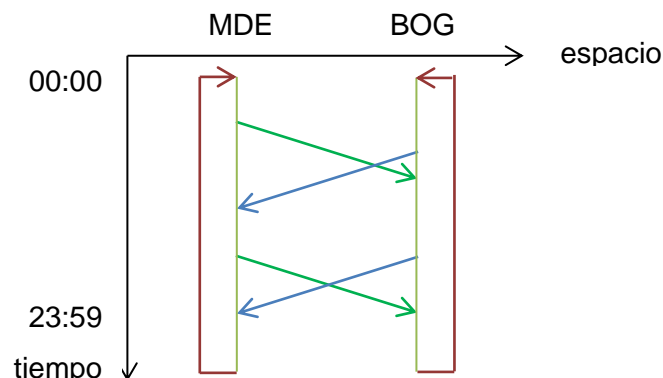
Red en el tiempo y en el espacio

Para plantear las restricciones se usa una red en el tiempo y en el espacio (*Time-space network*), común en todas las etapas de planeación en las que deben considerarse movimientos de entidades desde orígenes hacia destinos. Formalmente, una red en el tiempo y en el espacio para el problema integrado del Diseño de itinerarios y de la Asignación de Tipo de Flota consiste en superposición de un conjunto de redes, una para cada tipo de avión. En esta red, los nodos representan eventos en el tiempo y en el espacio (por ejemplo un despegue o un aterrizaje de un tipo de avión en cierto aeropuerto a cierta hora). Por su parte, hay tres tipos de arcos:

- Arcos en tierra: aviones que quedan en tierra entre dos eventos consecutivos.
- Arcos de vuelo: un vuelo entre un origen y un destino.
- Arcos de pernocta: conexión entre el primer y último nodo del día en cada aeropuerto.

Esta red permite al modelo elegir las conexiones entre vuelos, siempre y cuando sean factibles; es decir, siempre que se garantice un tiempo mínimo de conexión

Figura 1-5: Esquematización de una red en el tiempo y en el espacio



Elaboración propia con base en (Sherali et al., 2006)

1.2.5 Modelo matemático explícito para usar como base para proponer el cambio de resolución

El siguiente modelo, desarrollado por Sherali et al. (2006) y retomado en la propuesta de Abdelghany et al. (2017) se usará como base, en el Capítulo 3, para proponer la metodología de cambio de resolución aplicada al problema integrado de diseño de itinerarios y asignación de tipo de avión.

Conjuntos

S conjunto de aeropuertos de la red, indexados por s, o, d.

F conjunto de tipos de avión, indexado por f.

L conjunto de vuelos programables, indexados por l o {o, d, t}, donde o, d \in S y t es el tiempo de salida o llegada desde el origen o, o en el destino d.

N conjunto de nodos de la red de cada tipo de avión, indexados por {f, s, t}, f \in F, s \in S y t es un momento del tiempo

O(f) conjunto de arcos del tipo de avión f que cruzan la línea de conteo de aviones.

π conjunto de todos los itinerarios, indexados por i.

L^F: conjunto de vuelos obligatorios, indexados por l.

L^O: conjunto de vuelos opcionales (susceptibles de ser eliminados), indexados por l

π^O : conjunto de itinerarios que contienen vuelos opcionales (susceptibles de ser eliminados) indexados por i.

L(i) conjunto de vuelos en el itinerario i, i \in Π .

Parámetros

C_{fl}: costo de asignar el tipo de avión f al vuelo l.

A_f: número de aviones disponibles del tipo f

R_i: pérdida de rentabilidad si el itinerario i se excluye de la red.

t- y t+, los momentos previos y posterior al momento t en la línea del tiempo.

Variables de decisión

$$X_{fl} = \begin{cases} 1 & \text{si el avión de tipo } f \text{ se asigna al vuelo } l \\ 0 & \text{en caso contrario} \end{cases}$$

Esta variable también podría definirse como

$$X_{fodt} = \begin{cases} 1 & \text{si el avión de tipo } f \text{ se asigna para viajar de } o \text{ a } d \text{ en el tiempo } t \\ 0 & \text{en caso contrario} \end{cases}$$

$Y_{fstt'}$: flujo de aviones en tierra en el tipo de avión f en la estación s desde el nodo en tierra $\{f,s,t\}$ al nodo $\{f,s,t'\}$ en entre los momentos t y t' con $t' > t$.

$$Z_i = \begin{cases} 1 & \text{si el itinerario } i \text{ se incluye en el itinerario final} \\ 0 & \text{en caso contrario} \end{cases}$$

Función objetivo

$$\text{Min} \sum_{l \in L} \sum_{f \in F} C_{fl} X_{fl} + \sum_{i \in \pi^0} R_i (1 - Z_i) \quad (1.1)$$

En la ecuación (1.1), la primera sumatoria refleja el costo de asignación de tipo de avión. La segunda sumatoria refleja la pérdida de rentabilidad por excluir los itinerarios. Nótese que la expresión Maximizar (Ganancias – Costos) usada en la descripción verbal es equivalente a Minimizar (Costos + Ganancias perdidas) utilizada de forma algebraica en la ecuación (1.1).

Restricciones

Balance:

$$\sum_{o \in S} X_{fost} + Y_{fst-t} = \sum_{d \in S} X_{fsdt} + Y_{fstt'} \quad \forall \{f,s,t\} \in N \quad (1.2)$$

Recuérdese que cada nodo de la red en el tiempo y en el espacio representa un evento (por ejemplo un despegue o un aterrizaje). La ecuación de balance (1.2) garantiza en cada nodo que la cantidad de aviones que aterrizan más los aviones que había en tierra antes de dicho evento es igual a la cantidad de aviones que despegan más la cantidad de aviones que se quedan en tierra después del evento.

Disponibilidad:

$$\sum_{l \in O(f)} X_{fl} + \sum_{s \in S} Y_{fst_n t_1} \leq A_f \quad \forall f \in F \quad (1.3)$$

La ecuación de disponibilidad (1.3) garantiza que la cantidad de aviones que se usan de cada tipo no excede la cantidad disponible. Esta ecuación se plantea en un momento específico (y arbitrario) del horizonte de planeación (denominado count time en inglés), en el que debe cumplirse que la cantidad de aviones en vuelo más la cantidad de aviones en

tierra (de cada tipo) en dicho momento del tiempo no debe exceder la cantidad de aviones disponibles de cada tipo.

Cobertura de vuelos obligatorios:

$$\sum_{f \in F} X_{fl} = 1 \quad \forall l \in L^F \quad (1.4)$$

La ecuación (1.4) garantiza que los vuelos obligatorios queden efectivamente asignados a un único tipo de avión.

Cobertura de vuelos opcionales:

$$\sum_{f \in F} X_{fl} \leq 1 \quad \forall l \in F^O \quad (1.5)$$

La ecuación (1.5) verifica que, en caso de que un vuelo opcional sea elegido para formar parte del itinerario, se asigne a un único tipo de avión; en caso contrario, no se asigna a ningún tipo de avión.

Inclusión/exclusión de itinerarios:

$$z_i \leq \sum_{f \in F} X_{fl} \quad \forall i \in \pi^O, l \in L(i) \quad (1.6)$$

$$z_i - \sum_{i \in L(i)} \sum_{f \in F} X_{fl} \geq 1 - |L(i)| \quad \forall i \in \pi^O \quad (1.7)$$

La ecuación (1.6) indica que al excluir un itinerario completo, ninguno de sus vuelos se incluye en el itinerario final. La ecuación (1.7) verifica que si no todos los vuelos se asignan, el itinerario no se asigna.

Tipos de variables

$$X_{fl}, Z_i \in \{0,1\}. Y_{fstt'} \geq 0, Y_{fstt'} \in \mathbb{Z} \quad (1.8)$$

1.2.6 Preguntas abiertas en investigación en modelos de optimización para la planeación de la operación del transporte aéreo

En los trabajos analizados, el paradigma mayoritario es que partiendo de un itinerario base (diseñado manualmente o utilizado en el pasado) se proponen mejoras incrementales que consisten en añadir, eliminar o modificar las horas de salida de algunos vuelos, y se deja al modelo de optimización la elección de cuáles vuelos de la lista maestra elegir para el

itinerario final y a qué tipo de avión se asigna cada vuelo elegido. La lista maestra que se suministra tiene como datos fijos la hora de salida de cada vuelo, lo cual elimina la posibilidad de elegir (considerando cualquier criterio de eficiencia) la mejor hora de salida para cada vuelo. Unas pocas propuestas, como la de Abdelghany et al. (2017) consideran una ventana de tiempo para cada vuelo, pero que de todos modos limita la elección de cada hora de salida a un conjunto discreto. Esto deja ver que siguen abiertas cuestiones relacionadas con las formulaciones de los modelos, como la posibilidad de flexibilizar la hora de salida de los vuelos, de forma tal que la elección se haga en un espectro continuo (y por lo tanto sea posible usar programas lineales) buscando favorecer las conexiones de usuarios en hubs. Asimismo, un modelo que no tenga prefijada la cantidad de vuelos entre cada par origen-destino, sino que esta cantidad haga parte de las variables de decisión del problema, daría elementos para valorar, desde la optimización matemática, el intercambio (*trade-off*) entre la porción del mercado que la aerolínea desea capturar y la capacidad (en términos de cantidad de aviones) que debe tener para cumplir este objetivo. Por otro lado, y aunque en la literatura se evidencia un fuerte trabajo en el desarrollo de heurísticas que permitan solucionar rápidamente los modelos de optimización (en virtud de su carácter NP-hard), una pregunta que está abierta, y que es sobre la que se trabajará en esta tesis, tiene que ver con la interacción de los modelos de optimización con otras herramientas matemáticas que permitan hacer un manejo eficiente de los datos, y así verificar si una agrupación de datos reduce drásticamente el tiempo de solución del modelo, y más importante, si mejora la calidad de la solución encontrada.

2.Marco teórico sobre herramientas de Inteligencia Artificial para agrupar datos

2.1 Introducción

En el contexto de los datos, un **clúster** puede definirse como un grupo de observaciones que son similares entre ellas y disímiles de las observaciones de otros clústers, según las características de las observaciones (Yannakakis & Togelius, 2017). Dentro de las múltiples técnicas de Aprendizaje de Máquina e Inteligencia Artificial, la segmentación y la clusterización de datos son dos herramientas que permiten encontrar la estructura subyacente a los datos a partir de similitudes entre ellos. En general, los algoritmos de clusterización se usan para producir subgrupos homogéneos a partir de un conjunto de observaciones en las que sus clases o *etiquetas* no han sido identificadas previamente. Estas clasificaciones pueden ser no triviales respecto a las que se podrían resultar luego de aplicar criterios intuitivos.

Los algoritmos de clusterización utilizan medidas de distancia entre los datos con el fin de encontrar aquellos que son más cercanos entre sí (y que por lo tanto son similares). Como en todos los problemas en los que se cuenta con magnitudes medidas en distintas unidades, antes de pasar por un algoritmo de clusterización, las características de las observaciones deben estar en una escala común para tener distancias comparables. Dentro de las múltiples metodologías de normalización de datos, dos de las más usadas, por ejemplo en el análisis multicriterio y en el manejo de la distribución normal estándar, son las normalizaciones min-max, descrita en la ecuación (2.1) y z, descrita en la ecuación (2.2).

$$X'_{ij} = \frac{X_{ij} - \min_j [X_{ij}]}{\max_j [X_{ij}] - \min_j [X_{ij}]} \quad (2.1)$$

$$X'_{ij} = \frac{X_{ij} - \mu(X_{ij})}{\sigma(X_{ij})} \quad (2.2)$$

En ambas ecuaciones, X'_{ij} es el valor normalizado de la característica i de la observación j del conjunto de datos. En la ecuación (2.1), $\min_j [X_{ij}]$ corresponde al mínimo valor de la característica i de todas las observaciones y $\max_j [X_{ij}]$ corresponde al máximo valor de la característica i de todas las observaciones. En la ecuación (2.2), $\mu(X)$ es la media y $\sigma(X)$ es la desviación estándar de la característica i de las observaciones. Al usar la normalización min-max, $0 \leq X'_{ij} \leq 1$, mientras que si se usa la normalización z , se espera que la gran mayoría de normalizaciones queden en el intervalo $[-3,3]$, en virtud del comportamiento de la distribución normal.

Como se mencionó, el concepto de distancia es fundamental para establecer cuáles observaciones se encuentran más cerca entre sí, para poder segmentar o clusterizar los datos. Para calcular la distancia entre dos puntos p y q , pueden usarse muchas fórmulas, cada una con métricas distintas. En la ecuación (2.3) se presenta la fórmula de la distancia Manhattan (de métrica 1) y en la ecuación (2.4) se presenta la fórmula de la distancia Euclidiana (de métrica 2).

$$d_{PQ} = \sum_{i=1}^n |p_i - q_i| \quad (2.3)$$

$$d_{PQ} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (2.4)$$

En las ecuaciones (2.3) y (2.4), n corresponde a la cantidad de *dimensiones* de las observaciones que se están tratando. La principal diferencia entre las dos ecuaciones es que las diferencias más grandes entre p_i y q_i son penalizadas con más fuerza en la distancia euclidiana, en virtud del exponente 2. Una modificación de la ecuación (2.4) consistiría en usar métricas mayores a 2. A mayor métrica, se introduciría una penalización mayor a la heterogeneidad en las dimensiones. Así, una forma general de expresar la distancia entre dos puntos, dada una métrica q , se muestra en la ecuación (2.5).

$$d_{PQ} = \left(\sum_{i=1}^n (p_i - q_i)^q \right)^{\frac{1}{q}} \quad (2.5)$$

2.2 Características básicas del aprendizaje supervisado y no supervisado

Los métodos de aprendizaje de máquina pueden dividirse en dos grandes grupos: métodos de aprendizaje supervisado y métodos de aprendizaje no supervisado. En el primero, el insumo principal es un conjunto de datos, cada uno con una etiqueta o clasificación. El hecho de contar con una clasificación *a priori*, dada a través de un conjunto de datos de entrenamiento, reduce la tarea del algoritmo de clasificación a “únicamente” encontrar el patrón existente entre los datos que tienen cada etiqueta. Una vez se ha logrado este aprendizaje, se usa un conjunto de datos de validación para comparar su etiqueta real vs la que predice el algoritmo y calcular una medida de error. En el aprendizaje no supervisado, no existe una etiqueta asignada para cada dato y por lo tanto no hay manera de calcular una medida de error para un conjunto de datos de entrenamiento. La clusterización es una técnica de aprendizaje no supervisado ampliamente utilizada para agrupar ítems similares en clústers homogéneos en su interior y heterogéneos entre sí. La Tabla 2-1 muestra una breve síntesis de las características de cada tipo de aprendizaje.

Tabla 2-1: Características del aprendizaje supervisado y no supervisado

	¿Datos etiquetados a priori?	Objetivo del algoritmo	¿Se puede calcular una medida de error?	Tipos de algoritmos
Aprendizaje supervisado	Sí	Segmentar, asignar datos a etiquetas	Sí	K-NN Árboles de decisión Máquinas de soporte vectorial
Aprendizaje no supervisado	No	Agrupar elementos similares	No	K-Means DBSCAN

Elaboración propia con base en (Harrington, 2015; Namratha & Prajwala, 2012; Yannakakis & Togelius, 2017)

Algoritmos de k vecinos más cercanos (K-NN) y de clusterización (K-Means)

El algoritmo K-NN (k vecinos más cercanos), por sus siglas en inglés, es uno de los algoritmos de aprendizaje supervisado más ampliamente usados para clasificar datos en categorías. El uso de este algoritmo exige contar con un conjunto de datos de entrenamiento, en donde cada observación tiene una etiqueta o clasificación. La esencia

del algoritmo es usar las etiquetas ya conocidas para comparar observaciones de prueba (sin etiquetar todavía) con las del conjunto de entrenamiento, de manera que se le asigne a la observación sin clasificar la etiqueta que tengan la mayoría de sus k vecinos más cercanos (en términos de la distancia). El valor de k limita el tamaño del vecindario de cada observación de prueba y es elegido a priori. Al final, se compara la etiqueta predicha por el algoritmo vs la de la observación, y se calcula una medida de error (por ejemplo, el MSE). El pseudocódigo del algoritmo de clasificación k -NN se muestra a continuación:

INICIO

Leer: conjunto de datos de entrenamiento, conjunto de datos de prueba, k

Para cada observación de los datos de prueba:

 Calcular distancia a cada observación de los datos de entrenamiento

 Seleccionar las k observaciones de los datos de entrenamiento con menor distancia

 Votación: conteo de la clasificación de los k vecinos seleccionados

 Asignación: la observación de prueba se etiqueta con la clasificación con mayor votación

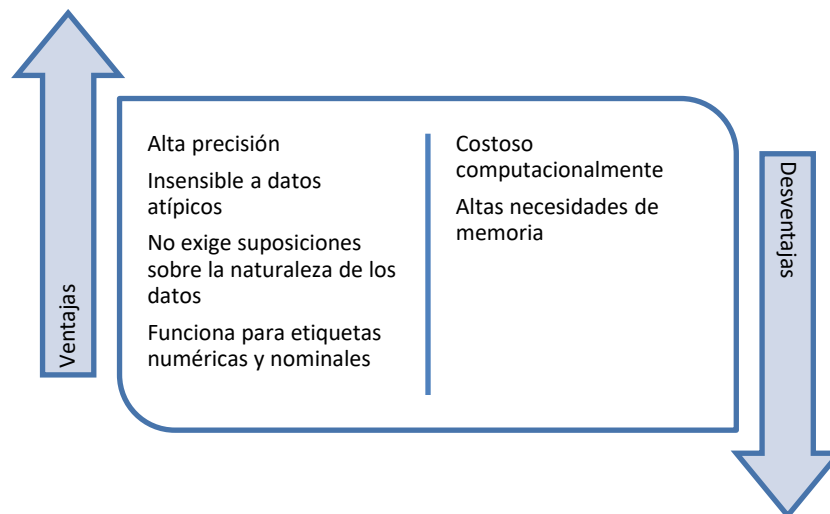
Fin para

 Calcular medida de error del algoritmo

FIN

En la Figura 2-1 se muestran algunas ventajas y desventajas de la clasificación k -NN.

Figura 2-1: Ventajas y desventajas de la clasificación k -NN



Fuente: (Harrington, 2015)

Por su parte, el algoritmo de clusterización k-Means es uno de los más usados para aprendizaje no supervisado. La esencia de este algoritmo consiste en definir inicialmente (por lo general de forma aleatoria) k centroides (en el dominio de las observaciones que se van a agrupar). Seguidamente, cada observación se asigna a un clúster al encontrar el centroide más cercano. Cuando se han formado los k clústers con las observaciones, el centroide de cada clúster se actualiza calculando la media aritmética de todos los puntos que pertenecen a él. Luego, cada punto se reasigna a cada clúster iterativamente hasta que los centroides no se actualicen (es decir, hasta que no haya ninguna reasignación de puntos a clústers) (Namratha & Prajwala, 2012). Como puede verse, la elección aleatoria de los centroides iniciales hace que la clusterización no obtenga la misma asignación de puntos a clústers en cada corrida del algoritmo, por lo cual el algoritmo corre en riesgo de caer en mínimos locales.

El pseudocódigo del algoritmo de clusterización k-Means se muestra a continuación:

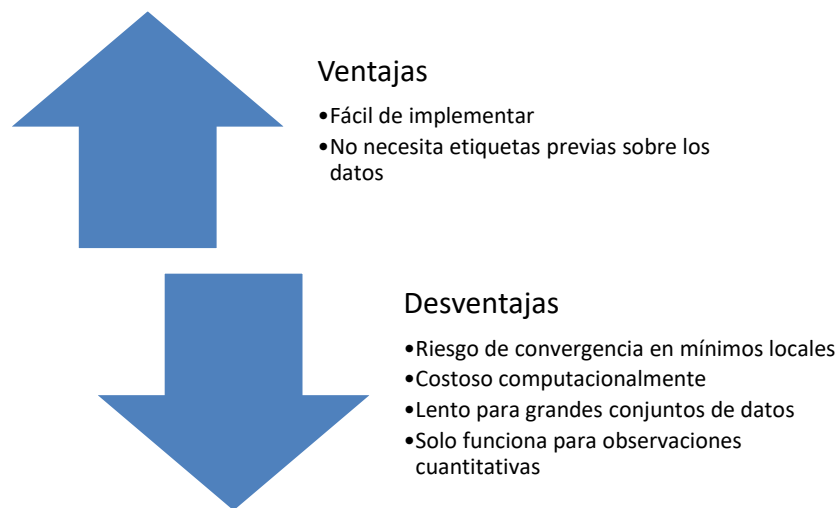
INICIO

```
Leer: conjunto de datos de prueba, k
Crear k centroides (generalmente de forma aleatoria)
Mientras haya cambios en la asignación de puntos a clústers:
    Para cada observación del conjunto de prueba:
        Para cada centroide:
            Calcular distancia entre observación y
            centroide
        Fin para
    Asignar el punto al clúster con menor distancia
    Fin para
    Para cada centroide:
        Calcular la media de las observaciones en el
        clúster
        Asignar el centroide a la media calculada
    Fin para
Repetir
```

FIN

En la Figura 2-2 se muestran algunas ventajas y desventajas de la clusterización k-Means.

Figura 2-2: Ventajas y desventajas de la clasificación k-NN



Fuente: (Harrington, 2015)

2.3 Aplicaciones generales de clusterización para reducir la cantidad de datos

Con la creciente capacidad de los computadores para generar, procesar y almacenar datos, mediciones y estadísticas, uno de los problemas que han venido enfrentando los tomadores de decisiones ha sido cómo manejar grandes volúmenes de datos y extraer la información más relevante de ellos. Desde el enfoque de bases de datos han surgido enfoques, programas y software orientados exclusivamente al Big Data, para manejar de forma sistemática grandes volúmenes de datos, provenientes de fuentes tan diversas como las redes sociales, sensores meteorológicos, sistemas de información financieros, entre otros.

Con tantos datos disponibles, muchas veces se hace imposible trabajar con todos ellos de forma individual: puede ser que al usar tal cantidad de datos se desborde la capacidad computacional de los sistemas soporte a la decisión (es decir, no de la base de datos sino del software que usa los datos para apoyar la toma de decisiones) o que, incluso teniendo una adecuada capacidad computacional, cada dato individual no agregue información significativa para el análisis. En este contexto, la segmentación o clusterización de datos

por categorías homogéneas en su interior y heterogéneas entre sí adquiere relevancia como un elemento de apoyo al procesamiento de datos previo al análisis de decisiones, tanto desde la perspectiva computacional como desde la interpretación física que resulta de agrupar elementos similares entre sí. Algunas aplicaciones generales de la clusterización de datos incluyen la división de clientes por áreas geográficas para entrega de correo, la agrupación de tipos de vegetaciones para estudios de hábitat, la clasificación de píxeles para reconocimiento de patrones, el diseño de sistemas de distribución y la asignación de clientes a diferentes segmentos de mercadeo de acuerdo con sus características demográficas (Bard & Jarrah, 2009). Como puede verse, la clusterización no necesariamente se hace con el propósito de aligerar la cantidad de parámetros de un modelo de optimización: hay casos en los que simplemente analizar dato a dato no es relevante y se hace necesaria una agregación en una escala temporal o espacial para obtener información útil.

El caso analizado por Bard y Jarrah (2009) es el de una compañía transportadora de carga que desea reducir sus costos fijos y para ello desea encontrar una configuración homogénea de áreas a servir por un mismo tipo de vehículo (recoger y entregar paquetes) cumpliendo restricciones operativas. En este caso, la segmentación de datos (la generación de áreas geográficas susceptibles de ser servidas por cada tipo de vehículo) se justifica ya que, de acuerdo con los autores, la cantidad de parámetros de cualquier instancia de prueba para el modelo de *picking* y *delivery* va desde 6.000 hasta 50.000 datos. Para agrupar las ubicaciones geográficas, el procedimiento desarrollado consistió en construir grupos de puntos geográficos (de recogida o de entrega de paquetes) de forma iterativa añadiendo secuencialmente puntos a los clústers hasta que se cumpliera a tope alguna de las restricciones operativas del sistema (por ejemplo, el tiempo de recorrido máximo permitido de un vehículo). Con un problema similar, de administración de una cadena de suministro, Kargari y Sepehri (2012) exponen el caso de una distribuidora de autopartes que debe categorizar a los clientes de un conjunto de almacenes según su comportamiento de compra, con el propósito de disminuir los costos de transporte de piezas a cada almacén. La agrupación de clientes se da ante la imposibilidad de estudiar el comportamiento de cada comprador individual para decidir el patrón de envío de partes a cada almacén.

En el estudio de las Redes Neuronales Artificiales, uno de los problemas comunes que se tratan es el de dividir el conjunto de datos con el que se trabaja, en subconjuntos de entrenamiento, validación y prueba. De acuerdo con el estudio de Sahoo, Zuo, y Tiwari (2012), la investigación en redes neuronales (y en mecanismos de inteligencia artificial en general) se ha centrado en su diseño e implementación y no tanto en el manejo de datos. Es por ello que, utilizando reglas heurísticas y redes neuronales de prueba, los autores dividieron los datos en los tres subconjuntos, cuidándose de mantener una proporción entre cada subconjunto de datos, y encontraron que la varianza del error obtenido por la red es más estable al usar la agrupación siguiendo reglas de clusterización, en comparación con una división trivial. En el artículo resaltan pruebas con proporciones 70%:20%:10% y 50%:40%:10%.

Otro caso que vale la pena analizar sobre clusterización de datos se da en biología: Milone, et al. (2013) describen el caso del uso de mapas auto-organizativos en el estudio de información de sistemas biológicos. En su artículo, los autores explican que en esta rama de la biología, donde cada objeto es descrito por una gran cantidad de variables (o dimensiones), es crucial realizar análisis exploratorios, integración de datos y descubrimiento de relaciones entre variables en un espacio *pequeño*, de manera que se pueda extraer la información relevante. Para este propósito, los mapas auto-organizativos (redes neuronales artificiales entrenadas para representar en pocas dimensiones los datos de entrada) resultan altamente útiles dada su capacidad de preservar la topología de los datos de entrada, por lo que reducen un espacio de búsqueda de numerosas dimensiones, en un mapa 2D que permite apreciar las agrupaciones y conexiones entre datos. En este caso, no se desea optimizar ninguna medida de desempeño, sino utilizar conexiones y agrupaciones de datos para entender las relaciones subyacentes entre ellos.

Por otro lado, en el mercado energético, se resaltan dos aplicaciones de la clusterización de datos: Härtel, Kristiansen, y Korpås (2017) exploraron y evaluaron las implicaciones de diferentes enfoques de reducción de dimensión para la planificación de la expansión de la transmisión energética a largo plazo al involucrar fuentes renovables de energía. Las razones que argumentaron para hacer esto es que la planeación de la expansión de la transmisión en el corto y largo plazo resulta muy compleja al involucrar fuentes renovables de energía, dada la variabilidad e incertidumbre sobre su disponibilidad. La agrupación de

datos en una escala temporal y espacial se hace con el fin de mantener la tratabilidad del problema de planeación. La estrategia que utilizaron para agrupar los datos consistió en normalizarlos según el valor más alto en cada categoría tecnológica analizada y comparar diferentes técnicas de clusterización (k-Means, K-Medoids, Moment-Matching y clusterización jerárquica). Encontraron que la agrupación de datos no necesariamente produce resultados confiables en el modelo de planeación de la expansión de la transmisión de energía y que debe trabajarse para mejorar la confiabilidad de los resultados. Por otro lado, Unternährer, Moret, Joost, y Maréchal (2017) exploraron el problema de clusterizar nodos de demanda de calefacción para ser suministrada por medio de distritos térmicos. En este problema, la cantidad de nodos de demanda que deben atenderse desborda la capacidad computacional del modelo de asignación que se desea correr, por lo que se optó por agruparlos usando la metodología k-Means. La conclusión principal de este artículo es que la agrupación de nodos de demanda para la planeación de la operación de distritos térmicos se ve promisorio como una estrategia que permita facilitar el análisis de demanda.

Finalmente, y aunque no se realiza una reducción en la cantidad de parámetros involucrada, vale la pena mencionar el trabajo de Sherali, Bae, & Haouari (2010), quienes plantearon un algoritmo para resolver el problema integrado de definición de itinerarios y asignación de flota usando Descomposición de Benders. Esta es una metodología de resolución de problemas de optimización en la que el problema original se divide en dos niveles (denominados maestro y subproblemas), cada uno con una porción de las variables de decisión, que de forma coordinada van asignando valores de las variables de decisión hasta llegar a la solución óptima del problema. En la aplicación de Sherali et al. (2010), las variables relacionadas con la inclusión de vuelos se asignan al problema maestro y las variables relacionadas con asignación de tipo de avión a los vuelos elegidos se deja al subproblema. Esta partición responde a las tipologías de variables de cada decisión exigidas en la Descomposición de Benders para el problema maestro y el subproblema, pero también puede hacerse una interpretación física, al separar las decisiones de cada etapa, para facilitar cálculos, pero coordinar esas decisiones para garantizar una decisión final factible y óptima. Con la estrategia de Descomposición de Benders, si bien se hace una segmentación de variables, no se reduce la cantidad total que se utiliza en el problema. Esa situación puede seguir comprometiendo la posibilidad de resolver un modelo de optimización en un solver corriente, y por eso es que para el caso de la planeación de la

operación de una aerolínea, se hace necesario pensar en estrategias de segmentación de datos que efectivamente reduzcan el tamaño del problema.

En cada uno de los anteriores ejemplos puede verse cómo la agrupación de datos por categorías o según su ubicación espacial es una estrategia usada para reducir el tamaño de los problemas a resolver, o para obtener información útil agregando conjuntos de datos, que individualmente no aportarían información que oriente adecuadamente la toma de decisiones.

3. Modelo propuesto de solución integrada de los problemas de diseño de itinerarios y de asignación de flota utilizando una estrategia de cambio de resolución

3.1 Introducción

Los modelos planteados por Lohatepanont (2002) y Sherali, Bae, y Haouari (2010) que integran las etapas de diseño de itinerarios y de asignación de flota carecen de flexibilidad para asignar las horas de salida de los vuelos. El modelo de Abdelghany, Abdelghany, y Azadian (2017) permite al modelo de optimización elegir la mejor hora de salida de cada vuelo (en un conjunto discreto de copias de un vuelo originalmente propuesto) pero prefija el tipo de avión que servirá cada vuelo. En este capítulo se propone un modelo de optimización para resolver de manera integrada los problemas de diseño de itinerarios y de asignación de flota, eligiendo la mejor hora de salida de cada vuelo (también en una ventana de tiempo alrededor de la hora de salida originalmente propuesta). Como se puede ver, el modelo propuesto busca capturar las mejores características de los modelos existentes: la resolución integrada de las etapas de diseño de itinerarios y de asignación de flota, permitiendo la elección de la hora de salida más conveniente. El costo de esta propuesta es, naturalmente, el aumento del tamaño combinatorial del modelo (en términos de la cantidad de variables y de ecuaciones necesarias para plantearlo, y del tiempo de solución requerido por el solver para entregar la solución óptima del problema). Por ello, también en este capítulo se propone una estrategia de clusterización de conjuntos de datos que utiliza herramientas de Aprendizaje de Máquina y de Inteligencia Artificial para reducir el tamaño combinatorial del modelo resultante, en términos de la cantidad de variables de decisión y de ecuaciones necesarias para resolver el problema.

3.2 Suposiciones

- Se cuenta con un itinerario base: un conjunto de vuelos que servirán a ciudades ya seleccionadas. Los vuelos están clasificados en dos conjuntos excluyentes y exhaustivos: vuelos obligatorios y vuelos opcionales.
- Se cuenta también con una flota definida, tanto en tipos de aviones como en cantidad de aviones de cada tipo.
- Los tiempos de vuelo entre cada par de ciudades no dependen del tipo de avión. Esta es una simplificación que se hace por practicidad.
- Para cada vuelo propuesto en el itinerario base se crearán a copias en una ventana de tiempo de longitud b creada alrededor de la hora de salida original.
- El itinerario base es para un día; el itinerario final resultante se repetirá todos los días de la semana. Esta es una simplificación que ignora la estacionalidad en la demanda de viajes entre días de semana y días de fin de semana, pero que se hace con fines de practicidad. Esta suposición es común en los procesos de planeación en las aerolíneas reales (Dong et al., 2016). Dado el caso, se podría crear un itinerario base para cada día de la semana.
- El costo de operar un vuelo en cada tipo de avión se calculará de forma proporcional a la distancia entre las ciudades servidas.
- La rentabilidad de cada vuelo se calculará (de forma simplificada) igualmente de forma proporcional a la distancia entre las ciudades servidas.
- Los costos y rentabilidades serán iguales para los vuelos originales y sus copias. Esto no necesariamente es cierto, porque la demanda de un vuelo puede ser sensible a la hora de salida.

3.2.1 Entradas del modelo

- Itinerario base (conjunto de vuelos, cada uno con origen, destino, hora de salida y hora de llegada).
- Tipos de aviones (y cantidad de aviones de cada tipo) disponibles.
- Número de copias que se realizarán de cada vuelo del itinerario base.
- Longitud de la ventana de tiempo.

3.2.2 Salidas del modelo

- Itinerario final propuesto: vuelos elegidos para ser servidos en un tipo de avión determinado y saliendo a una hora específica.
- Número de conexiones exitosas entre los vuelos incluidos en el itinerario final. Se dice que una conexión es exitosa cuando el avión o tripulación del vuelo de llegada puede ser usada en un vuelo de salida más tarde. Una conexión infactible es una conexión no exitosa entre dos vuelos.
- Costo total del itinerario final propuesto.

3.3 Formulación algebraica del modelo propuesto

En esta sección se formula de manera algebraica el modelo propuesto para resolver de forma integrada los problemas de diseño de itinerarios y asignación de flota, permitiendo la elección de la hora de salida de cada uno de los vuelos.

3.3.1 Conjuntos *iniciales*

S: conjunto de aeropuertos (estaciones) en la red: El conjunto S contiene todos los aeropuertos (estaciones) en los que ocurrirán vuelos en el itinerario base.

F: conjunto de tipos de avión: El conjunto F contiene únicamente los tipos de avión. Más adelante, como parámetro, se indicará la cantidad disponible de aviones de cada tipo. Los tipos de aviones difieren en sus características técnicas (cantidad de asientos, costos operativos, rango de alcance) y por lo tanto presentarán costos y rentabilidades diferentes para los vuelos entre diferentes pares de ciudades.

L: conjunto original de vuelos programables: Resultado de un análisis previo o con base en un itinerario pasado, se cuenta con un itinerario base, representado por un conjunto de vuelos. Cada vuelo está caracterizado por su número de vuelo, origen, destino, hora de salida y hora de llegada. Como se mencionó, el conjunto L contiene vuelos para un día, que serán replicados los demás días de la semana. En adelante, cuando se mencionen los elementos del conjunto L, se hará referencia a los números de vuelo.

L^F : conjunto de vuelos obligatorios: L^F es un subconjunto de F que contiene los vuelos que de manera obligatoria deben incluirse en el itinerario final. Esta decisión puede basarse en que dichos vuelos se han comportado satisfactoriamente en el pasado, tanto desde la perspectiva financiera como operacional. En el problema de optimización, la decisión asociada a los vuelos obligatorios tiene que ver únicamente con la asignación del tipo de avión que cubrirá ese vuelo y la hora de salida (entre el conjunto discreto de posibilidades).

L^O : conjunto de vuelos opcionales (susceptibles de ser eliminados): El conjunto de vuelos opcionales corresponde a los vuelos que no necesariamente deben quedar incluidos en el itinerario final. Estos vuelos pueden corresponder a modificaciones de vuelos de una temporada pasada, o a vuelos que deseen añadirse por primera vez al itinerario. En el problema de optimización, las decisiones asociadas a los vuelos opcionales tienen que ver con su inclusión en el itinerario final, y, en caso afirmativo, el tipo de avión asignado y la hora de salida (entre el conjunto discreto de posibilidades).

Note que $L = L^F \cup L^O$; $L^F \cap L^O = \emptyset$.

3.3.2 Parámetros para crear copias de cada vuelo

El conjunto L de vuelos originales, que es el que debe proponerse de manera inicial, no es el conjunto de vuelos que se someterá a la evaluación del modelo de optimización: de cada vuelo original se crearán copias, con horas de salida distribuidas en una ventana tiempo alrededor de la hora de salida propuesta para el vuelo original.

Sean:

a: número de copias que se crearán de cada vuelo

b: longitud de la ventana de tiempo

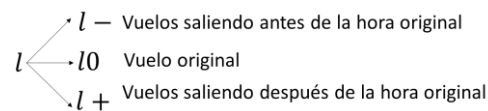
Así, cada vuelo original será representado por $a + 1$ vuelos. Si a es un número par y se desea que las copias de los vuelos sean equiespaciadas, la hora de salida (o de llegada) del j -ésimo vuelo copia (HS_j) puede calcularse, a partir de la hora de salida (o de llegada) del vuelo original (HS) como se ve en la ecuación (3.1).

$$HS_j = HS + j \frac{b}{a}, \quad j = -\frac{a}{2}, \dots, \frac{a}{2} \quad (3.1)$$

3.3.3 Notación de los vuelos copia

El vuelo original l tiene asociados $a + 1$ vuelos en el conjunto *ampliado* de vuelos que se ingresará al modelo de optimización. Al número del vuelo se le añadirá un sufijo: negativo para las copias que salen antes, 0 para representar el vuelo original, y positiva para las copias que salen más tarde, según se esquematiza en la Figura 3-1.

Figura 3-1: Esquematización de la notación de los vuelos copia



Elaboración propia

3.3.4 Otros conjuntos

Conjunto ampliado de vuelos: Sea V el conjunto ampliado de vuelos. Este conjunto tiene los vuelos originales y copias, con la respectiva notación y con las horas de salida y llegadas calculadas usando la ecuación (3.1).

Conjunto de vuelos en el aire a la medianoche: Una vez se tiene el conjunto V , es necesario crear un subconjunto V^R , que contiene todos los vuelos que a la medianoche estarían en vuelo; es decir, aquellos vuelos que despegarían un día y aterrizarían al siguiente (en caso de ser incluidos en el itinerario). Este subconjunto es importante porque en las ecuaciones de balance calculadas para el último nodo de cada aeropuerto), no se tendría la totalidad de los aviones en tierra, y por tanto el conteo de aviones debe considerar también la cantidad de aviones en vuelo.

N: conjunto de nodos de la red en el tiempo y en el espacio de cada tipo de avión: Un nodo es un evento: una salida o una llegada de un vuelo a un aeropuerto a una hora determinada. Por tanto, cada vuelo tiene asociados dos nodos: uno de salida (en el aeropuerto de origen) y uno de llegada (en el aeropuerto de destino). Los nodos se nombrarán así, donde l es el número del vuelo:

- l_{out} : nodos de salida.
- l_{in} : nodos de llegada.

Cada aeropuerto tiene un subconjunto de nodos, ordenados cronológicamente.

$U(N)$: conjunto de últimos nodos del día en cada aeropuerto: En las ecuaciones de balance que se formularán más adelante, se garantizará el flujo de aviones a través de la red en el tiempo y en el espacio. Para ello se debe considerar con especial interés los últimos nodos de cada aeropuerto (según el orden cronológico de ocurrencia) y agruparlos en un subconjunto U .

$P(N)$: conjunto de primeros nodos del día en cada aeropuerto: Por razones análogas a las que se enuncian para formar el conjunto U , se debe crear el subconjunto P .

3.3.5 Parámetros

C_{fv} : Costo de asignar el tipo de avión f al vuelo v : es un costo *operativo* que depende de forma general del tipo de avión asignado y de la distancia entre las ciudades servidas.

A_f : Número de aviones disponibles del tipo f : en el parámetro A_f se condensa la disponibilidad de flota con la que cuenta la aerolínea.

R_i : Pérdida de rentabilidad si el vuelo i se excluye del itinerario final: al igual que el costo operativo, la pérdida de rentabilidad por no incluir un vuelo en el itinerario final se calculará de forma proporcional a la distancia entre dos ciudades. Esta simplificación no es completamente realista, dado que la rentabilidad depende también de la demanda en el par origen-destino. La pérdida de rentabilidad se calcula para los vuelos originales y no para los vuelos copia. Esto se hace con el propósito de contar una sola vez la exclusión de un vuelo opcional del itinerario (y no la exclusión de todas las copias de ese vuelo).

$Conect_{vn}$: Una matriz de conectividad entre vuelos y nodos. Este arreglo permitirá involucrar, en las ecuaciones de balance, los vuelos entrantes y salientes de cada nodo. La definición de los elementos de la matriz se muestra en la ecuación (3.2)

$$Conect_{vn} = \begin{cases} 1 & \text{si el vuelo } v \text{ es una llegada al nodo } n \\ -1 & \text{si el vuelo } v \text{ es una salida del nodo } n \\ 0 & \text{en caso contrario} \end{cases} \quad (3.2)$$

K_{lv} : Una matriz de asociación entre vuelos originales y vuelos copias. Este arreglo permitirá relacionar, en las ecuaciones de cubrimiento de vuelos, únicamente las copias de cada vuelo original. La definición de los elementos de la matriz se muestra en la ecuación (3.3).

$$K_{lv} = \begin{cases} 1 & \text{si el vuelo } v \text{ es una copia del vuelo } l \\ 0 & \text{en caso contrario} \end{cases} \quad (3.3)$$

HS_v : Hora de salida del vuelo v . Este parámetro se crea con base en la tabla de vuelos copia y únicamente se extrae la hora de salida.

HL_v : Hora de llegada del vuelo v . Este parámetro también se crea con base en la tabla de vuelos copia.

$Salida_l$: Aeropuerto de salida del vuelo l . Este parámetro se extrae de la tabla de vuelos originales.

$Llegada_l$: Aeropuerto de llegada del vuelo l . Este parámetro se extrae de la tabla de vuelos originales.

ρ : El tiempo mínimo que garantiza una conexión exitosa entre dos vuelos (entrante y saliente) en un aeropuerto.

3.3.6 Variables de decisión

Como se ha mencionado, el interés del modelo propuesto es decidir de manera simultánea si un vuelo es elegido para formar parte del itinerario final, y el tipo de avión que cubrirá dicho vuelo. En ese sentido se define una variable binaria X_{fv} que tomará el valor de 1 si el avión de tipo f se asigna al vuelo v , y 0 en caso contrario. Por otro lado, para garantizar el balance de aviones de cada tipo, se debe crear una variable entera Y_{fn} que permita llevar la cuenta del número de aviones de tipo f que se encuentran en tierra en el nodo n . Finalmente, y con el propósito de privilegiar los vuelos con conexiones exitosas entre sí, se crea una variable Z_{ij} , que tomará el valor de 1 si existe una conexión exitosa entre

alguna de las copias de los vuelos originales i y j . Así, las variables de decisión se muestran en la ecuación (3.4).

$$X_{fv} = \begin{cases} 1 & \text{si el avión de tipo } f \text{ se asigna al vuelo } v \\ 0 & \text{en caso contrario} \end{cases}$$

$$Y_{fn}: \text{número de aviones de tipo } f \text{ en tierra en el nodo } n \quad (3.4)$$

$$Z_{ij} = \begin{cases} 1 & \text{si hay conexión exitosa entre las copias de los vuelos originales } i \text{ y } j \\ 0 & \text{en caso contrario} \end{cases}$$

3.3.7 Función objetivo

La función objetivo del modelo se muestra en la ecuación (3.5).

$$\text{Min} \sum_f \sum_v C_{fv} X_{fv} + \sum_l r_l \left(1 - \sum_f \sum_v K_{lv} X_{fv} \right) + M_1 \sum_{(i,j) \text{ tales que } \text{Salida}_j = \text{Llegada}_i} 1 - Z_{ij} \quad (3.5)$$

La primera sumatoria de la ecuación (3.5) calcula el costo de asignación de los vuelos a los tipos de aviones. Note que, en virtud de la variable binaria X_{fv} , únicamente se tienen en cuenta los vuelos elegidos en el itinerario final. La segunda sumatoria, por su parte, calcula la pérdida de rentabilidad asociada a la exclusión de vuelos del itinerario. De nuevo, como se tiene una variable binaria X_{fv} , únicamente se cuentan las pérdidas de rentabilidad de los vuelos excluidos del itinerario final. Finalmente, la tercera sumatoria castiga la inclusión de vuelos con conexiones infactibles entre sí. M_1 es un valor con el que se castigan las conexiones infactibles. Para solucionar el modelo, M_1 debe tener un valor lo suficientemente severo para castigar los pares de vuelos con conexiones infactibles, pero no tan alto como para que se sacrifique la rentabilidad de la operación.

3.3.8 Restricciones

Balance:

$$Y_{fn} = Y_{f,n-1} + \sum_v \text{Conect}_{vn} X_{fv} \quad \forall f \in F \quad \forall n \notin P(N) \quad (3.6)$$

$$Y_{fn} = Y_{fu} + \sum_v \text{Conect}_{vn} X_{fv} \quad \forall f \in F \quad \forall n \in P(N) \quad (3.7)$$

En las ecuaciones (3.6) y (3.7) se calcula el balance de aviones de cada tipo en cada nodo. La restricción asegura que el número de aviones de cada tipo en cada nodo (en un instante de tiempo) equivale a la cantidad de aviones que había en el nodo anterior (en el instante de tiempo anterior), más la cantidad de aviones entrantes (vuelos que llegan a ese nodo, con $Conect_{vn} = 1$ en ese nodo), menos la cantidad de aviones salientes (vuelos que salen de ese nodo, con $Conect_{vn} = -1$ en ese nodo). La ecuación (3.6) se construye para todos los nodos con excepción de los primeros de cada aeropuerto, en cuyo caso se usa la ecuación (3.7). En dicha ecuación, el nodo u es el que, en el conjunto $U(N)$ ocupa la misma posición del nodo n en el conjunto $P(N)$.

Disponibilidad:

$$\sum_{n \in U} Y_{fn} + \sum_{v \in V^R} X_{fv} \leq A_f \quad \forall f \in F \quad (3.8)$$

La ecuación (3.8) se construye para garantizar que no se usen más aviones que los disponibles de cada tipo. La manera de hacerlo es sumando los aviones de cada tipo que quedan en tierra en el último nodo del día en cada aeropuerto (en los nodos del conjunto U) con la suma de los vuelos en el aire a la medianoche (en caso de que sean incluidos en el itinerario) y verificar que esta cantidad de aviones sea menor o igual que el valor del parámetro A_f .

Cobertura de vuelos obligatorios:

$$\sum_f \sum_v K_{lv} X_{fv} = 1 \quad \forall l \in L^F \quad (3.9)$$

En la ecuación (3.9) se garantiza, para cada uno de los vuelos originales marcados como obligatorios, que se asignen a uno de los tipos de avión disponible, y que únicamente se elija una de las copias de dicho vuelo.

Cobertura de vuelos opcionales:

$$\sum_f \sum_v K_{lv} X_{fv} \leq 1 \quad \forall l \in L^O \quad (3.10)$$

En la ecuación (3.10) se garantiza, para cada uno de los vuelos originales marcados como opcionales, que, en caso de ser elegidos para formar parte del itinerario final, se asignen

a uno de los tipos de avión disponible, y que únicamente se elija una de las copias de dicho vuelo.

Relación entre las variables X y Z:

$$Z_{ij} \leq \frac{(\sum_f \sum_v K_{fv} HS_v X_{fv} - \sum_f \sum_v K_{iv} HL_v X_{fv} - \rho)}{M_2} + 1 \quad \forall i \forall j \in L \mid Salida_j = Llegada_i \quad (3.11)$$

$$2Z_{ij} \leq \sum_f \sum_v K_{iv} HL_v X_{fv} + \sum_f \sum_v K_{jv} HS_v X_{fv} \quad \forall i \forall j \in L \mid Salida_j = Llegada_i \quad (3.12)$$

$$Z_{ij} = 0 \text{ siempre que } Salida_j \neq Llegada_i \quad (3.13)$$

La ecuación (3.11) garantiza que la variable Z_{ij} toma el valor de 1 cuando el tiempo de conexión entre las copias elegidas de los vuelos entrante i y saliente j es superior a ρ . En dicha ecuación, M_2 es un valor muy grande respecto a los datos del problema. La ecuación (3.12) garantiza que la variable Z_{ij} únicamente puede tomar el valor de 1 cuando tanto el vuelo entrante y el vuelo saliente j son incluidos en el itinerario. La ecuación (3.13) fija con el valor de 0 la variable Z_{ij} cuando el vuelo entrante i y el vuelo saliente j no ocurren en el mismo aeropuerto.

Factibilidad de la operación

$$X_{fv} = 0 \text{ siempre que el avión de tipo } f \text{ no pueda cubrir el vuelo } v \quad (3.14)$$

La ecuación (3.14) garantiza que únicamente se cubrirán vuelos en tipos de avión *apropiados*. Esto depende de factores como la distancia que deben viajar los aviones entre las ciudades servidas (que debe ser menor que el rango máximo de vuelo del tipo de avión) y la longitud de pista disponible para despegar y aterrizar (que debe ser mayor que la longitud requerida por el tipo de avión para el despegue y aterrizaje).

Tipos de variables:

$$X_{fv}, Z_{ij} \in \{0,1\} \quad \forall f \in F, \forall v \in V, \forall i \forall j \in L \quad (3.15)$$

$$Y_{fn} \geq 0, Y_{fn} \in \mathbb{Z} \quad \forall f \in F \forall n \in N \quad (3.16)$$

En la ecuación (3.15) se definen X_{fv} y Z_{ij} como variables binarias y en la ecuación (3.16) se asegura que las variables Y_{fn} tomen valores enteros no negativos.

3.4 Planteamiento de un caso básico de estudio

Con el propósito de explicar la estructura y razón de existencia de cada elemento del modelo propuesto, se desarrollará un caso básico de estudio.

3.4.1 Datos del caso básico de estudio

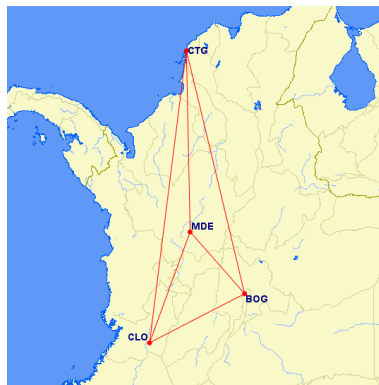
Conjuntos *iniciales*

S: conjunto de aeropuertos (estaciones) en la red: Se cuenta con cuatro aeropuertos: Medellín, Bogotá, Cali y Cartagena, identificados con el código IATA de sus aeropuertos y mostrados como conjunto en la ecuación (3.17).

$$S = \{MDE, BOG, CLO, CTG\} \quad (3.17)$$

En la Figura 3-2 se muestran los cuatro aeropuertos. Las líneas que unen cada par de ciudades indican que habrá vuelos entre ellas.

Figura 3-2: Aeropuertos considerados en el caso básico de estudio



Fuente: Elaboración propia usando Great Circle Mapper

F: conjunto de tipos de avión: Se cuenta con dos tipos de aviones: Airbus A320 (A320) y Boeing 757 (B757), presentados como conjunto en la ecuación (3.18).

$$F = \{A320, B757\} \quad (3.18)$$

Figura 3-3: Tipos de aviones disponibles en el caso básico de estudio



Imágenes obtenidas en línea

L: conjunto original de vuelos programables: Se considerarán 36 vuelos en total, cubriendo así todas las posibles combinaciones entre pares de ciudades. Los vuelos se muestran en la Tabla 3-1. Las horas de salida y llegada están expresadas en formato militar; así, por ejemplo, las 6 representa las 6:00 am y las 18 representa las 6:00 pm.

Tabla 3-1: Vuelos originales del caso básico de estudio

Número de vuelo	Origen	Destino	Salida	Llegada
001	MDE	BOG	6	7
002	MDE	BOG	10	11
003	MDE	BOG	19	20
004	BOG	MDE	7	8
005	BOG	MDE	13	14
006	BOG	MDE	16	17
007	MDE	CTG	9	10
008	MDE	CTG	13	14
009	MDE	CTG	16	17
010	CTG	MDE	11	12
011	CTG	MDE	15	16
012	CTG	MDE	21	22
013	MDE	CLO	7	8
014	MDE	CLO	12	13
015	MDE	CLO	15	16
016	CLO	MDE	8	9
017	CLO	MDE	10	11
018	CLO	MDE	18	19
019	BOG	CTG	5	6
020	BOG	CTG	11	12
021	BOG	CTG	18	19
022	CTG	BOG	10	11
023	CTG	BOG	13	14
024	CTG	BOG	19	20
025	BOG	CLO	6	7
026	BOG	CLO	10	11
027	BOG	CLO	17	18
028	CLO	BOG	6	7
029	CLO	BOG	10	11
030	CLO	BOG	19	20
031	CLO	CTG	8	9
032	CLO	CTG	12	13
033	CLO	CTG	15	16
034	CTG	CLO	10	11
035	CTG	CLO	14	15
036	CTG	CLO	15	16

Fuente: Elaboración propia

Como puede verse en la Tabla 3-1, cada vuelo de los 36 vuelos propuestos tiene un origen, destino, hora de salida y hora de llegada.

L^F : conjunto de vuelos obligatorios: El conjunto de vuelos obligatorios contiene 28 de los 36 vuelos originales, según se muestra en la ecuación (3.19).

$$L^F = \{001, 002, 003, 006, 007, 008, 009, 011, 012, 013, 014, 016, 017, 018, 019, 020, 021, 022, 024, 025, 026, 027, 029, 030, 031, 032, 035, 036\} \quad (3.19)$$

L^O : conjunto de vuelos opcionales (susceptibles de ser eliminados): Los vuelos opcionales son los 8 vuelos no marcados como obligatorios de los 36 originales. Se muestran en la ecuación (3.20).

$$L^O = \{004, 005, 010, 015, 023, 028, 033, 034\} \quad (3.20)$$

Como puede verse, $L = L^F \cup L^O$; $L^F \cap L^O = \emptyset$

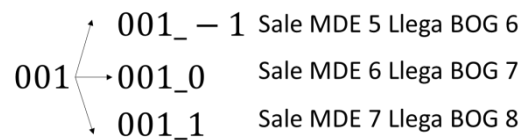
Parámetros para crear copias de cada vuelo

Para el caso básico de estudio, se supondrá $a = 2$; $b = 2$. Esto implica que cada vuelo tendrá 3 versiones: la original, una copia 1 (saliendo y llegando una hora antes de las horas originales) y una copia 2 (saliendo y llegando una hora después de las horas originales).

Notación de los vuelos copia

Siguiendo la notación descrita, se nombrarán todos los vuelos copia de este caso básico. Por ejemplo, las copias del vuelo 001 serán nombradas en el conjunto ampliado de vuelos como se indica en la Figura 3-4.

Figura 3-4: Ejemplo de la notación de vuelos copia en el caso básico de estudio



Elaboración propia

En la Tabla 3-2 se muestra el conjunto V para el caso básico de estudio.

Tabla 3-2 Conjunto ampliado de vuelos para el caso básico de estudio

Número de vuelo	Origen	Destino	Salida	Llegada
001_-1	MDE	BOG	5	6
001_0	MDE	BOG	6	7
001_1	MDE	BOG	7	8
002_-1	MDE	BOG	9	10
002_0	MDE	BOG	10	11
002_1	MDE	BOG	11	12
003_-1	MDE	BOG	18	19
003_0	MDE	BOG	19	20
003_1	MDE	BOG	20	21
004_-1	BOG	MDE	6	7
004_0	BOG	MDE	7	8
004_1	BOG	MDE	8	9
005_-1	BOG	MDE	12	13
005_0	BOG	MDE	13	14
005_1	BOG	MDE	14	15
006_-1	BOG	MDE	15	16
006_0	BOG	MDE	16	17
006_1	BOG	MDE	17	18
007_-1	MDE	CTG	8	9
007_0	MDE	CTG	9	10
007_1	MDE	CTG	10	11
008_-1	MDE	CTG	12	13
008_0	MDE	CTG	13	14
008_1	MDE	CTG	14	15
009_-1	MDE	CTG	15	16
009_0	MDE	CTG	16	17
009_1	MDE	CTG	17	18
010_-1	CTG	MDE	10	11
010_0	CTG	MDE	11	12
010_1	CTG	MDE	12	13
011_-1	CTG	MDE	14	15
011_0	CTG	MDE	15	16
011_1	CTG	MDE	16	17
012_-1	CTG	MDE	20	21
012_0	CTG	MDE	21	22

012_1	CTG	MDE	22	23
013_-1	MDE	CLO	6	7
013_0	MDE	CLO	7	8
013_1	MDE	CLO	8	9
014_-1	MDE	CLO	11	12
014_0	MDE	CLO	12	13
014_1	MDE	CLO	13	14
015_-1	MDE	CLO	14	15
015_0	MDE	CLO	15	16
015_1	MDE	CLO	16	17
016_-1	CLO	MDE	7	8
016_0	CLO	MDE	8	9
016_1	CLO	MDE	9	10
017_-1	CLO	MDE	9	10
017_0	CLO	MDE	10	11
017_1	CLO	MDE	11	12
018_-1	CLO	MDE	17	18
018_0	CLO	MDE	18	19
018_1	CLO	MDE	19	20
019_-1	BOG	CTG	4	5
019_0	BOG	CTG	5	6
019_1	BOG	CTG	6	7
020_-1	BOG	CTG	10	11
020_0	BOG	CTG	11	12
020_1	BOG	CTG	12	13
021_-1	BOG	CTG	17	18
021_0	BOG	CTG	18	19
021_1	BOG	CTG	19	20
022_-1	CTG	BOG	9	10
022_0	CTG	BOG	10	11
022_1	CTG	BOG	11	12
023_-1	CTG	BOG	12	13
023_0	CTG	BOG	13	14
023_1	CTG	BOG	14	15
024_-1	CTG	BOG	18	19
024_0	CTG	BOG	19	20
024_1	CTG	BOG	20	21
025_-1	BOG	CLO	5	6
025_0	BOG	CLO	6	7

025_1	BOG	CLO	7	8
026_-1	BOG	CLO	9	10
026_0	BOG	CLO	10	11
026_1	BOG	CLO	11	12
027_-1	BOG	CLO	16	17
027_0	BOG	CLO	17	18
027_1	BOG	CLO	18	19
028_-1	CLO	BOG	5	6
028_0	CLO	BOG	6	7
028_1	CLO	BOG	7	8
029_-1	CLO	BOG	9	10
029_0	CLO	BOG	10	11
029_1	CLO	BOG	11	12
030_-1	CLO	BOG	18	19
030_0	CLO	BOG	19	20
030_1	CLO	BOG	20	21
031_-1	CLO	CTG	7	8

031_0	CLO	CTG	8	9
031_1	CLO	CTG	9	10
032_-1	CLO	CTG	11	12
032_0	CLO	CTG	12	13
032_1	CLO	CTG	13	14
033_-1	CLO	CTG	14	15
033_0	CLO	CTG	15	16
033_1	CLO	CTG	16	17
034_-1	CTG	CLO	9	10
034_0	CTG	CLO	10	11
034_1	CTG	CLO	11	12
035_-1	CTG	CLO	13	14
035_0	CTG	CLO	14	15
035_1	CTG	CLO	15	16
036_-1	CTG	CLO	14	15
036_0	CTG	CLO	15	16
036_1	CTG	CLO	16	17

Fuente: Elaboración propia

Note que como había 36 vuelos originales, el conjunto ampliado de vuelos tiene $108 = 36 \times 3$ vuelos.

Otros conjuntos

Conjunto de vuelos en el aire a la medianoche: En este caso básico de estudio, no hay vuelos en el aire a la medianoche, por lo que $V^R = \emptyset$

N: conjunto de nodos de la red en el tiempo y en el espacio de cada tipo de avión: Como el conjunto V del caso básico de estudio tiene 108 vuelos, el conjunto N tendrá 216 nodos, esquematizados en la ecuación (3.21). Por ejemplo, el vuelo 001_-1 tiene asociados dos nodos: el nodo 001_-1_out, que se da en Medellín a las 5:00 am, y el nodo 001_-1_in, que ocurre en Bogotá a las 6:00 am.

$$N = \{001_-1_out, 001_-1_in, \dots, 036_1_out, 036_1_in\} \quad (3.21)$$

Cada aeropuerto tiene un subconjunto de nodos, ordenados cronológicamente.

U(N): conjunto de últimos nodos del día en cada aeropuerto: Para el caso básico de estudio, U cuenta con cuatro elementos, uno para cada estación (MDE, BOG, CLO y CTG), que se muestran en la ecuación (3.22).

$$U = \{012_1_in, 030_1_in, 030_1_out, 012_1_out\} \quad (3.22)$$

P(N): conjunto de primeros nodos del día en cada aeropuerto: Para el caso básico de estudio, P está conformado por los nodos 001_-1_out (MDE), 019_-1_out (BOG), 028_-1_out (CLO), 019_-1_in (CTG).

Parámetros

C_{fv} : costo de asignar el tipo de avión f al vuelo v

Sea d_{od} la distancia entre las ciudades que se cubrirán por el vuelo v en el avión de tipo f entre las ciudades o y d . En el caso básico de estudio se asumirá el costo según la ecuación (3.23).

$$C_{fv} = \begin{cases} d_{od} & \text{si } f = A320 \\ 1.2d_{od} & \text{si } f = B757 \end{cases} \quad (3.23)$$

Las distancias entre ciudades se muestran en la Tabla 3-3.

Tabla 3-3: Distancias entre las ciudades consideradas en el caso básico de estudio

Ciudad 1	Ciudad 2	Distancia (mi)
MDE	BOG	134
MDE	CTG	295
MDE	CLO	192
BOG	CTG	406
BOG	CLO	175
CLO	CTG	478

Fuente: Elaboración propia usando Great Circle Mapper

Así, se obtiene un arreglo de costos según lo muestra la Tabla 3-4.

Tabla 3-4 Costos operativos calculados para el caso básico de estudio

Tipo de avión / Vuelo	001_-1	001_0	001_1	...	036_-1	036_0	036_1
A320	134	134	134	...	478	478	478
B757	161	161	161	...	574	574	574

Fuente: Elaboración propia

A_f : número de aviones disponibles del tipo f : Se supondrá una cantidad de aviones según la ecuación (3.24).

$$A_f = \frac{A320}{B757} \begin{bmatrix} 6 \\ 4 \end{bmatrix} \quad (3.24)$$

R_i : pérdida de rentabilidad si el vuelo i se excluye del itinerario final: Para calcular la pérdida de rentabilidad por excluir un vuelo de forma proporcional a la distancia entre las ciudades consideradas, se usará un factor de 1.25, según se aprecia en la ecuación (3.25). La elección de este factor se hizo para garantizar que las rentabilidades de los vuelos sean mayores que los costos (que fueron calculados usando factores de 1 y 1.2).

$$r_i = 1.25d_{od} \quad (3.25)$$

De esta forma se obtiene un arreglo como el que se muestra en la Tabla 3-5.

Tabla 3-5: Rentabilidades calculadas para el caso básico de estudio

Vuelo	Pérdida de rentabilidad
001	168
002	168
003	168
...	...
034	598
035	598
036	598

Fuente: Elaboración propia

HS_v : Hora de salida del vuelo v : El parámetro HS_v se muestra en la Tabla 3-6.

Tabla 3-6 Hora de salida de cada vuelo en el caso básico de estudio

v	HS_v
001_-1	5
001_0	6
...	...
036_1	16

Elaboración propia

HL_v : Hora de llegada del vuelo v : El parámetro HL_v es se muestra en la Tabla 3-7.

Tabla 3-7 Hora de llegada de cada vuelo en el caso básico de estudio

v	HS_v
001_-1	6
001_0	7
...	...
036_1	17

Elaboración propia

Salida_i: Aeropuerto de salida del vuelo i . El vector se muestra en la Tabla 3-8.

Tabla 3-8 Aeropuerto de salida de cada vuelo en el caso básico de estudio

v	Salida _v
001	MDE
002	MDE
...	...
036	CTG

Elaboración propia

Llegada: El parámetro Llegada_i se muestra en la Tabla 3-9.

Tabla 3-9 Aeropuerto de llegada de cada vuelo en el caso básico de estudio

v	Salida _v
001	BOG
002	BOG
...	...
036	CLO

Elaboración propia

ρ : El tiempo mínimo que garantiza una conexión exitosa entre dos vuelos (entrante y saliente) en un aeropuerto. Se asumirá un valor de media hora para el parámetro ρ , según se ve en la ecuación (3.26).

$$\rho = 0.5 \quad (3.26)$$

Función objetivo

Para la sumatoria que penaliza la inclusión de vuelos con conexiones infactibles entre sí, se asumió un valor de $M_1 = 125$.

Factibilidad de la operación

Ninguna variable X_{fv} se prefijó como 0, lo que indica que todos los vuelos pueden ser asignados a cualquier tipo de avión.

3.4.2 Programación del caso básico de estudio

El modelo de optimización para resolver el caso básico de estudio fue programado en GAMS 24.7.3. El código del programa se encuentra en el Anexo A y el archivo .gms se puede encontrar en el siguiente [vínculo](#).

3.4.3 Resultados del caso básico de estudio

En este modelo propuesto interesa saber cuáles vuelos opcionales fueron incluidos en el itinerario, a qué tipo de avión fue asignado cada uno de los vuelos (tanto obligatorios como opcionales), cuál copia de cada vuelo original fue elegida, cuántos aviones de cada tipo se utilizaron, cuántas conexiones exitosas fueron asignadas, y cuál es el costo del itinerario sugerido por el modelo de optimización. En la Tabla 3-10 se muestran los resultados.

Tabla 3-10: Resultados del caso básico de estudio

Vuelo	Copia elegida	Tipo de avión elegido
001	001_-1	B757
002	002_0	A320
003	003_1	A320
004	004_0	A320
005	005_-1	A320
006	006_-1	B757
007	007_0	A320
008	008_-1	A320
009	009_1	A320
010	010_-1	A320
011	011_-1	A320
012	012_0	A320
013	013_-1	A320
014	014_1	A320
015	015_1	A320
016	016_-1	A320
017	017_1	A320
018	018_0	A320

019	019_1	A320
020	020_1	A320
021	021_-1	B757
022	022_-1	A320
023	023_1	A320
024	024_1	A320
025	025_-1	A320
026	026_-1	A320
027	027_-1	A320
028	028_0	A320
029	029_0	B757
030	030_-1	A320
031	031_-1	A320
032	032_0	A320
033	033_1	A320
034	034_1	B757
035	035_0	A320
036	036_-1	A320

Fuente: Elaboración propia

El costo del itinerario elegido para el caso básico de estudio es de \$32.596 unidades monetarias imaginarias (incluyendo costos de operación, pérdidas de rentabilidad e inclusión de conexiones infactibles). Este itinerario incluye 146 conexiones exitosas y el uso de 6 A320 (de 6 disponibles) y 3 Boeing 757 (de 4 disponibles).

3.4.4 Tamaño del modelo

En la Tabla 3-11 se resumen algunas características del modelo, en términos de la cantidad de variables y de ecuaciones, y el tiempo de procesamiento y solución del modelo, usando los datos del caso básico de estudio. El modelo se resolvió utilizando la plataforma Neos Server, que usó un servidor del Departamento de Ciencias de la Computación de la Universidad de Wisconsin-Madison. Los servidores de este departamento cuentan con procesadores de 2.2, 2.3 y 2.8 GHz, y memorias RAM de 64 y 192 GB. Las características detalladas de los servidores pueden consultarse en el siguiente [vínculo](#).

Tabla 3-11 Estadísticas del modelo usando los datos del caso básico de estudio

Número de variables de decisión	975
Número de ecuaciones (Función objetivo y restricciones)	1.121
Número de iteraciones requeridas para llegar a la solución óptima	881
Tiempo de solución requerido por el servidor disponible	0,014 s

De la Tabla 3-11 se destaca el hecho de que, para únicamente 36 vuelos originales, se hayan necesitado 975 variables de decisión (binarias y enteras) y 1121 ecuaciones.

3.5 Algoritmo propuesto de clusterización de datos

En esta sección se muestra cómo, usando el algoritmo *k-Means* de clusterización de datos explicado en el Capítulo 2, se logra formular el mismo problema de optimización del modelo propuesto, sometiendo uno de los conjuntos de datos a un proceso de clusterización, de manera que, al reducir su tamaño, se reduzca también el tamaño del problema de optimización (tanto en la cantidad de variables de decisión y ecuaciones como en el tiempo de solución). Este algoritmo se implementará para el caso de aplicación desarrollado en el Capítulo 4, ya que en el caso básico de estudio, como solo hay dos tipos de aviones, no se justifica realizar la clusterización.

3.5.1 Clusterización del conjunto F de tipos de avión

Se propone agrupar los tipos de avión en clústers, de manera que los aviones más similares entre sí (de acuerdo con sus características técnicas) constituyan un único elemento del conjunto F. Esto, en términos prácticos, conduce a que el proceso de decisión consista en asignar a cada vuelo incluido en el itinerario un tipo **genérico** de avión, en

lugar de un tipo específico. Agrupar los tipos de aviones en clústers reducirá el tamaño del conjunto F , y por lo tanto, si se corre el problema de optimización con la versión clusterizada de F , disminuirá el número de variables de decisión y de restricciones del problema, y por consiguiente el solver podrá encontrar la solución óptima en menor tiempo. La desventaja de esta clusterización es que la respuesta que entregue el modelo estará en una escala *gruesa* y por tanto no se contará con una asignación específica de un tipo de avión a cada vuelo incluido en el itinerario. En consecuencia, una vez se tengan las asignaciones preliminares de cada vuelo a cada tipo genérico de aviones, es necesario desagregar esta respuesta para obtener una solución final en la escala original del problema y que satisfaga todas las restricciones operativas del modelo inicial (balance, disponibilidad y cubrimiento de vuelos).

Entradas del algoritmo de clusterización

- Tipos de aviones originales
- Características técnicas de cada tipo de avión
- Número de clústers (k)

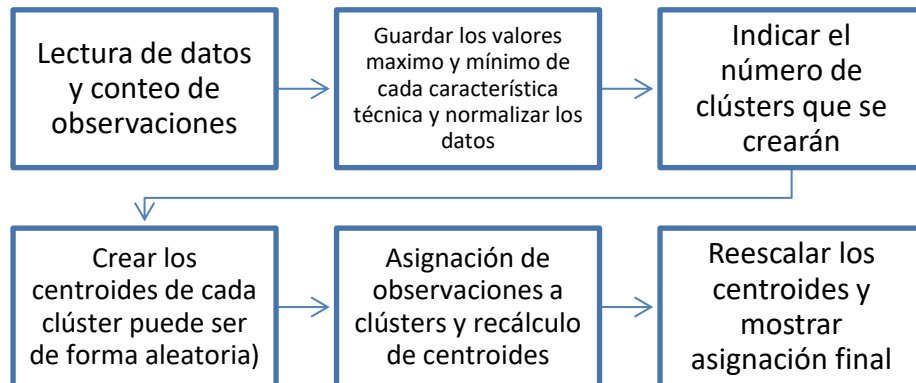
Salidas del algoritmo de clusterización

- Asignación de cada tipo de avión a uno de los k clústers
- Centroides de los clústers (expresados en la escala original de los datos)

3.5.2 Algoritmo de clusterización

En la Figura 3-5 se muestra una síntesis de la secuencia de pasos que se siguen para obtener cada clúster.

Figura 3-5: Síntesis de la secuencia de pasos para clusterizar conjuntos de datos



Elaboración propia

El algoritmo de clusterización propuesto que involucra la estrategia *k-Means* para agrupar los tipos de avión en clústers según sus similitudes técnicas se muestra en forma de pseudocódigo a continuación:

INICIO

```
Leer características técnicas de los tipos de aviones
Leer número de clústers que se crearán (k)
Crear k centroides (puede ser de forma aleatoria)
Mientras haya cambios en la asignación de tipos de aviones a clústers:
    Para cada tipo de avión:
        Para cada centroide:
            Calcular distancia entre tipo de avión y centroide (utilizando métrica 2)
        Fin para
        Asignar el tipo de avión al clúster con menor distancia
    Fin para
    Para cada centroide:
        Calcular la media de las observaciones en el clúster
        Asignar el centroide a la media calculada
    Fin para
Repetir
```

FIN

3.6 Adaptación del modelo propuesto a la clusterización del conjunto F

A continuación se muestra cómo el modelo propuesto debe modificarse ligeramente para incluir la clusterización de tipos de aviones.

3.6.1 Cálculo de parámetros que dependen del tipo de avión

Los parámetros C_{fv} (costo de asignar el tipo de avión f al vuelo v), A_f (número de aviones disponibles del tipo f) y W_{fl} (1 si el avión de tipo f no está habilitado para cubrir el vuelo l y 0 en caso contrario) dependen del tipo de avión. Cuando se plantea el modelo sin clusterizar datos, los parámetros técnicos de cada tipo de avión se pueden usar de forma directa en cada una de las ecuaciones usadas para definir los parámetros del modelo. En este caso, que ya no se tienen los tipos de avión de manera desagregada sino clusterizados, los parámetros técnicos que se usarán corresponderán a los que se obtengan en los centroides de los clústers creados.

3.6.2 Otros conjuntos y parámetros que no se alteraron debido a la clusterización

Los conjuntos S (aeropuertos), L (vuelos originales programables), L^F (vuelos obligatorios), L^O (vuelos opcionales), L^R (vuelos en el aire a la medianoche), V (conjunto ampliado de vuelos), N (nodos de la red), U (últimos nodos de cada estación) y P (primeros nodos de cada estación) se mantienen iguales respecto a la versión sin clusterizar debido a que estos conjuntos no dependen del conjunto F . En este mismo sentido, los parámetros a (número de copias que se crearán de cada vuelo), b (longitud de la ventana de tiempo), R_i (pérdida de rentabilidad por excluir un vuelo del itinerario), d_i (demanda de cada vuelo), $Conect_{vn}$ (matriz de asociación entre nodos y vuelos), K_{lv} (matriz de asociación entre vuelos originales y copias), HS_v (hora de salida de los vuelos), HL_v (hora de llegada de los vuelos), $Salida_i$ (aeropuerto de salida de los vuelos), $Llegada_i$ (aeropuerto de llegada de los vuelos) y ρ (tiempo mínimo para garantizar una conexión exitosa) se calculan de la misma forma que en el problema sin clusterización.

3.7 Estrategia de reconversión de resultados para asignar a cada vuelo elegido en el itinerario un tipo específico de avión y planteamiento de un modelo reducido de asignación

3.7.1 Formulación de la estrategia de reconversión de resultados

Cada uno de los vuelos incluidos en el itinerario se asignará a uno de los k clústers de tipos de aviones. Sin embargo, esta no es la respuesta que se debe entregar como solución del problema integrado de diseño de itinerarios y asignación de tipo de avión, debido a que las entradas del problema en cuanto a los tipos de aviones se expresaron de manera individual, y hasta este punto lo que se tiene es una asignación “preliminar” agregada de un clúster de aviones que se asignará al vuelo. La pregunta subyacente en este punto es: *¿cuál tipo de avión específico del clúster asignado será elegido para cubrir cada vuelo?* La estrategia elegida para “desclusterizar” las asignaciones obtenidas por el modelo consiste en correr, para cada clúster, una versión reducida del modelo de optimización general propuesto, en donde únicamente se consideran las restricciones de balance, disponibilidad, cobertura obligatoria y factibilidad para los vuelos asignados a cada clúster. En la Tabla 3-12 se resume la forma como se obtienen y calculan los conjuntos y parámetros del modelo de optimización reducido para cada clúster.

Tabla 3-12: Conjuntos y parámetros esenciales del modelo reducido de optimización para desclusterizar las respuestas

Conjunto / Parámetro	Obtención / Cálculo
F (Tipos de avión)	Se consideran únicamente los tipos de avión pertenecientes al clúster analizado.
L (Vuelos)	Los vuelos seleccionados en el itinerario y asignados al clúster analizado. Todos se consideran obligatorios.
N (nodos de la red)	Los nodos de salida y llegada de los vuelos asignados al clúster analizado.
C_{fi} (costo de asignar el tipo de avión f al vuelo i)	Calculado usando el CASM real de los tipos de avión del clúster analizado.
A_f (número de aviones disponibles del tipo f)	Únicamente se consideran los tipos de avión del clúster analizado.

Elaboración propia

En el modelo reducido se sigue manteniendo la estructura de las variables de decisión X y Y , pero no se incluyen las variables Z ni los parámetros involucrados para determinar si una conexión es exitosa o no. En la Tabla 3-13 se muestran los elementos del modelo

clusterizado que se obtienen o calculan igual para este modelo reducido, o que se modificaron o que se excluyeron para reducir el tamaño del modelo reducido.

Tabla 3-13: Otros conjuntos y parámetros del modelo reducido de optimización para desclusterizar las respuestas

Conjunto / Parámetro	Considerado igual / Modificado / Excluido del modelo reducido
L^F (subconjunto de vuelos obligatorios)	Excluido (todos los vuelos se consideraron obligatorios).
L^O (subconjunto de vuelos opcionales)	Excluido (todos los vuelos se consideraron obligatorios).
V (conjunto ampliado de vuelos (originales y copias))	Excluido (no se consideraron copias de los vuelos).
V^R (conjunto de vuelos en el aire a la medianoche)	Se cambia por L^R .
$U(N)$ (conjunto de últimos nodos del día en cada estación)	Considerado igual, únicamente para los aeropuertos asociados a los vuelos seleccionados en el itinerario y pertenecientes al clúster analizado.
$P(N)$ (conjunto de primeros nodos del día en cada estación)	Considerado igual, únicamente para los aeropuertos asociados a los vuelos seleccionados en el itinerario y pertenecientes al clúster analizado.
R_l (pérdida de rentabilidad si el vuelo l se excluye del itinerario)	Excluido. Este modelo reducido ya no elige los vuelos que se incluirán en el itinerario final
$Conect_{in}$ (matriz de conectividad entre vuelos y nodos)	Calculada igual que en el modelo con clusterización.
K_{iv} (matriz de asociación entre vuelos originales y copias)	Excluida. Este modelo reducido no considera vuelos copia.
HS_v (hora de salida del vuelo v)	Excluida. Este modelo reducido no considera conexiones entre vuelos.
HL_v (hora de llegada del vuelo v)	Excluida. Este modelo reducido no considera conexiones entre vuelos.
$Salida_l$ (aeropuerto de salida del vuelo l)	Excluido. Este modelo reducido no considera conexiones entre vuelos.
$Llegada_l$ (aeropuerto de llegada del vuelo l)	Excluido. Este modelo reducido no considera conexiones entre vuelos.
ρ (tiempo mínimo para que una conexión entre dos vuelos sea factible)	Excluido. Este modelo reducido no considera conexiones entre vuelos.
W_{fl} (matriz de infactibilidad de la operación del vuelo l con el tipo de avión f)	Calculada con los parámetros técnicos reales de los tipos de aviones pertenecientes al clúster analizado.

Elaboración propia

La disminución de la cantidad de vuelos y la eliminación del componente de conexiones (por haberse considerado ya en la optimización clusterizada) hacen que la optimización reducida sea esencialmente rápida y permita desagregar de forma factible los resultados obtenidos en la optimización clusterizada. La solución óptima de cada modelo reducido indicará la asignación de un tipo de avión perteneciente al clúster a cada uno de los vuelos asignados a él.

3.7.2 Formulación algebraica del modelo reducido de optimización

Variables de decisión

$$X_{fl} = \begin{cases} 1 & \text{si el tipo de avión } f \text{ del clúster analizado se asigna al vuelo } l \\ 0 & \text{en caso contrario} \end{cases}$$

$$Y_{fn}: \text{número de aviones en tierra del tipo} \quad (3.27)$$

$$\text{de avión } f \text{ del clúster analizado en el nodo } n$$

Función objetivo

$$\text{Min} \sum_f \sum_l C_{fl} X_{fl} \quad (3.28)$$

Restricciones

Balance:

$$Y_{fn} = Y_{f,n-1} + \sum_l \text{Conect}_{ln} X_{fl} \quad \forall f \in F \quad \forall n \notin P(N) \quad (3.29)$$

$$Y_{fn} = Y_{fu} + \sum_l \text{Conect}_{ln} X_{fl} \quad \forall f \in F \quad \forall n \in P(N) \quad (3.30)$$

Disponibilidad:

$$\sum_{n \in U} Y_{fn} + \sum_{l \in L^R} X_{fl} \leq A_f \quad \forall f \in F \quad (3.31)$$

Cobertura de vuelos:

$$\sum_f X_{fl} = 1 \quad \forall l \in L \quad (3.32)$$

Factibilidad de la operación:

$$X_{fl} = 0 \text{ cuando el vuelo } l \text{ no se pueda cubrir con el tipo de avión } f \quad (3.33)$$

Tipos de variables:

$$X_{fl} \in \{0,1\} \forall f \in F, \forall l \in L \quad (3.34)$$

$$Y_{fn} \geq 0, Y_{fn} \in \mathbb{Z} \forall f \in F \forall n \in N \quad (3.35)$$

El modelo reducido guarda semejanzas con el modelo clusterizado y con el modelo sin clusterización. Sin embargo, la disminución del tamaño de los conjuntos y la consideración de menos ecuaciones permiten un tiempo de solución y una cantidad de iteraciones considerablemente menor, según se mostrará con el caso de aplicación desarrollado en el Capítulo 4. Así, se logra obtener la solución del problema en su escala original, puesto que con la solución del modelo reducido para cada clúster se obtiene un tipo de avión específico asignado de manera factible a cada uno de los vuelos incluidos en el itinerario. Otra forma posible de realizar la clusterización podría ser agrupando los aeropuertos (por ejemplo según sus coordenadas geográficas), y la agregación se haría de forma semejante a como se hizo para los tipos de avión.

4.Caso de aplicación del modelo de optimización y de la estrategia de clusterización propuesta

4.1 Introducción

En este capítulo se presenta un caso de aplicación del modelo de optimización para resolver de manera integrada el problema de diseño de itinerarios y de asignación de flota usando el algoritmo de clusterización de datos propuesto. Se hace una comparación entre el tamaño del problema cuando no se usa clusterización vs cuando sí se usa, y se contrastan las soluciones obtenidas.

4.2 Descripción general del caso de aplicación

El caso que se tratará es una adaptación del itinerario real del otoño de 2010 de la Aerolínea Avianca. El propósito de usar un itinerario real es poder acceder a un conjunto de vuelos previamente concebido y con una asignación de flota hecha, para comparar las horas de salida y la asignación de tipos de avión a cada uno de los vuelos en el itinerario real y según los resultados del modelo. En la Figura 4-1 se puede apreciar un recorte del itinerario real. El archivo, al que se accedió de manera libre a través de la página web de la aerolínea y que se puede descargar usando el siguiente [vínculo](#) muestra, para cada par origen-destino, todos los vuelos programados con sus respectivas horas de salida, llegada, tipo de avión asignado, y una descripción de si los vuelos tienen o no escala, así como la operación de cada vuelo durante la semana. Cada día se representa por un número entre 1 y 7 (lunes, 1; martes, 2, y así sucesivamente). Cuando aparece el número, quiere decir que el vuelo se opera ese día. Cuando aparece un punto en la posición en la que debería estar un número, quiere decir que el vuelo no se opera ese día.

Figura 4-1: Recorte del itinerario de la aerolínea Avianca del otoño de 2010

# Vuelo	Hora Salida	Hora Llegada	Equipo (s)	P	Via	Frec	# Vuelo	Hora Salida	Hora Llegada	Equipo (s)	P	Via	Frec
SAN ANDRES ISLAND, GUSTAVO ROJAS PINILLA (ADZ)													
Hacia Bogotá							Desde Bogotá						
AV9787	12:05	15:30	MM100	1	CTG7	AV9786	08:10	11:40	MM100	1	CTG7
AV8491	13:20	16:45	MM100	1	CTG	...4.6.	AV8490	09:25	12:55	MM100	1	CTG	...4.6.
AV9779	14:05	16:05	AV320	0		1234567	AV9778	11:30	13:35	AV320	0		1234567
AV9777	16:50	18:50	AV320	0		1234567	AV9776	14:10	16:15	AV320	0		1234567
AV9781	22:00	00:10 (1)	MM100	0		1234567	AV9780	19:15	21:33	MM100	0		1234567

Fuente: Avianca

4.3 Suposiciones, adaptaciones y simplificaciones del itinerario para la formulación de caso de aplicación

- Para formar el itinerario base (utilizando como insumo el itinerario real) se supondrá que los vuelos obligatorios son los internacionales, y el 80% de los vuelos entre Bogotá y Medellín, Cali, Cartagena y Barranquilla, y que los vuelos opcionales son el resto de vuelos nacionales.
- Los tipos de aviones y la cantidad disponible de cada uno serán los que tenía la aerolínea en el 2010. Esta información es pública y se puede consultar en el Informe de Gestión de la empresa publicado en 2011, y en sitios libres de internet.
- Las horas de salida y llegada de los vuelos serán las publicados en el itinerario.
- Para cada vuelo se supondrá una demanda promedio de pasajeros.
- Para cada vuelo propuesto en el itinerario base, se crearán dos copias, en una ventana de tiempo de una hora creada alrededor de la hora de salida original.
- El itinerario base se asumirá para un día; el itinerario final resultante se repetirá todos los días de la semana. Si bien en el itinerario real se precisa cuáles vuelos se operarán cada día de la semana, en este caso de aplicación se asumirá la misma operación diaria, dado que el foco de la investigación está puesto en el efecto de la clusterización de datos para reducir el tamaño del problema. Sin embargo, si fuera el caso, podría plantearse un modelo para cada día de la semana por separado de forma semejante a como se plantea aquí.
- El costo de asignación de cada tipo de avión a cada vuelo tendrá dos componentes: un costo operativo y un costo de oportunidad que serán proporcionales a la distancia entre las ciudades servidas y a la cantidad de sillas del avión: para el costo operativo (asociado a la elección del tipo de avión), el factor de

proporcionalidad que se usará es el CASM (*cost per available seat mile*) para cada tipo de avión según promedios de la industria (Bazargan, 2010). El costo de oportunidad será derivado de ofrecer un tipo de avión con menos sillas que la demanda promedio del vuelo, y depende de una rentabilidad unitaria (RASM, *revenue per available seat mile*), de la distancia entre las ciudades servidas, y de la diferencia entre la demanda y el número de sillas ofrecido. El cálculo del costo operativo se muestra en la ecuación (4.1), el del costo de oportunidad en la ecuación (4.2) y el costo de asignación en la ecuación (4.3).

Sean:

d : distancia

s : número de sillas

$$C_{operativo} = CASM \times d \times s \quad (4.1)$$

$$C_{oportunidad} = \begin{cases} RASM \times distancia \times (d - s) & \text{si } d > s \\ 0 & \text{si } d \leq s \end{cases} \quad (4.2)$$

$$C_{asignación} = C_{operativo} + C_{oportunidad} \quad (4.3)$$

- La rentabilidad de cada vuelo se calculará (de forma simplificada) promediando el costo de asignación en cada tipo de avión (para dicho vuelo) y multiplicando ese resultado por el margen de utilidad de la industria, que puede ser tomado como un 5% (Belobaba et al., 2009), según se ve en la ecuación (4.4).

$$Rentabilidad = \overline{Costo} \times 1.05 \quad (4.4)$$

- Los costos y rentabilidades serán iguales para los vuelos originales y sus copias. De nuevo, y al igual que en el caso básico de estudio, esto no necesariamente es cierto, porque la demanda de un vuelo puede ser sensible a la hora de salida, pero esta simplificación es práctica dado el interés en comparar el tamaño del problema y el tiempo de solución cuando se clusterizan los datos vs cuando no se clusterizan.

4.4 Formulación algebraica del caso de aplicación sin involucrar clusterización de datos

Siguiendo la estructura de conjuntos, variables de decisión, función objetivo y restricciones presentadas en el Capítulo 3, aquí se presentan las fórmulas y datos para calcular los parámetros del modelo de este caso de aplicación.

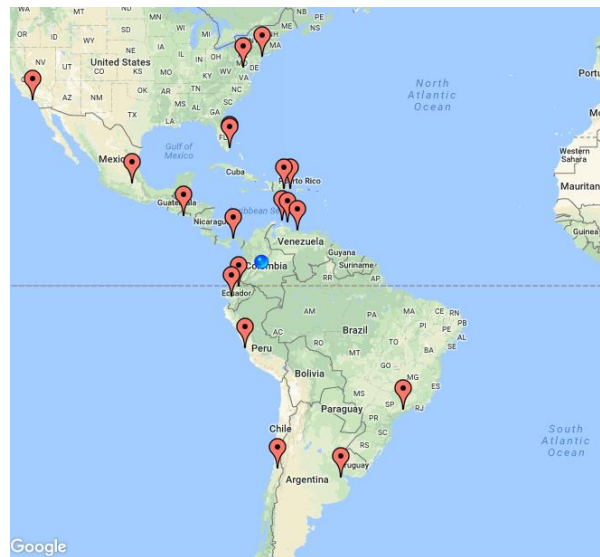
Conjuntos *iniciales*

S: conjunto de aeropuertos (estaciones) en la red: La red de rutas del caso de aplicación cuenta con 42 aeropuertos: 21 dentro de Colombia y 21 internacionales. Los aeropuertos se identificarán con sus códigos IATA. Los destinos nacionales se muestran en la Figura 4-2 y los destinos internacionales en la Figura 4-3, y los aeropuertos se muestran como conjunto en la ecuación (4.5).

Figura 4-2: Destinos nacionales del caso de aplicación



Figura 4-3: Destinos internacionales del caso de aplicación

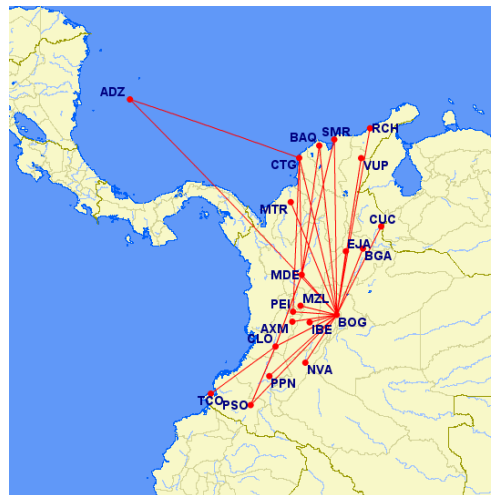


Fuente: Elaboración propia usando Google Maps

$$S = \left\{ \begin{array}{l} ADZ, AUA, AXM, BAQ, BCN, BGA, BOG, CCS, CLO, \\ CTG, CUC, CUR, EJA, EZE, FLL, GRU, GYE, IAD, \\ IBE, JFK, LAX, LIM, MAD, MDE, MEX, MIA, MTR, \\ MZL, NVA, PEI, PPN, PSO, PTY, PUJ, RCH, \\ SAL, SCL, SDQ, SMR, TCO, UIO, VUP \end{array} \right\} \quad (4.5)$$

En la Figura 4-4 se muestran las rutas nacionales planteadas y en la Figura 4-5 y Figura 4-6 las rutas internacionales.

Figura 4-4: Rutas consideradas en el caso de aplicación



Fuente: Elaboración propia usando Great Circle Mapper

Figura 4-5: Rutas internacionales a Centroamérica, Norteamérica y El Caribe consideradas en el caso de aplicación

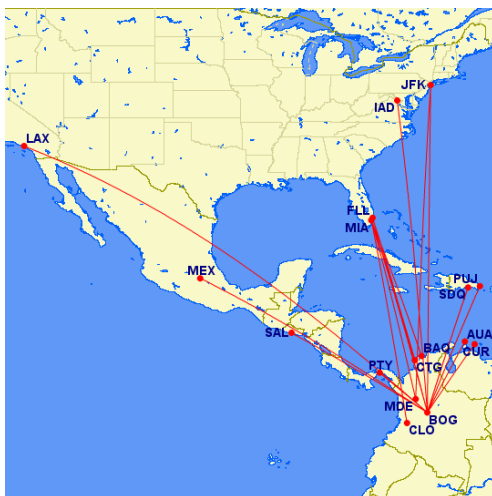
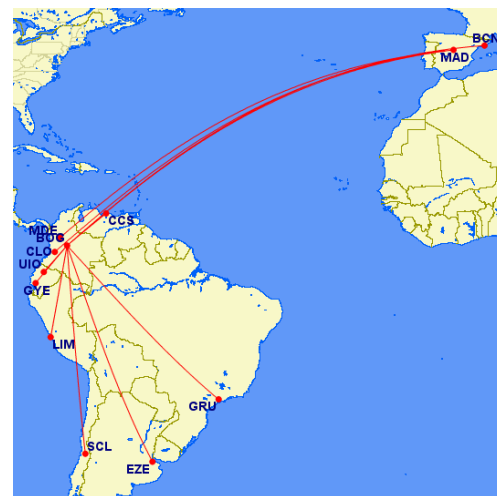


Figura 4-6: Rutas internacionales a Suramérica y Europa consideradas en el caso de aplicación



Fuente: Elaboración propia usando Great Circle Mapper

F: conjunto de tipos de avión: Para este caso de aplicación, los tipos de aviones son los siguientes:

- Airbus A330 (A330)
- Boeing 767 (B767)
- Airbus A320 (A320)
- Airbus A319 (A319)
- Airbus A318 (A318)
- Fokker 50 (F50)

Los tipos de aviones se presentan como conjunto en la ecuación (4.6).

$$F = \{A330, B767, A320, A319, A318, F50\} \quad (4.6)$$

Los datos técnicos de los aviones que se usarán para calcular los parámetros del modelo y como insumos para construir los clústers de datos, son: CASM (dólares), número de sillas, costo operativo por hora, rango de alcance (millas), la longitud de pista requerida para despegue (m) y longitud de pista requerida para aterrizaje (m). Estos datos se extrajeron del *Course Packet* de la asignatura *Airline Operations*, dictada en Purdue University y se muestran, para cada tipo de avión, en la Tabla 4-1.

Tabla 4-1: Datos técnicos de los aviones que se usarán para la clusterización

Tipo de avión	CASM (dólares)	Número de sillas	Costo operativo por hora (dólares)	Alcance (millas)	Longitud de pista requerida para despegue (m)	Longitud de pista requerida para aterrizaje (m)
A330	0,0465	280	8236	8333	2220	1750
B767	0,0612	278	9224	7587	2620	1524
A320	0,0627	150	4393	3509	1700	1463
A319	0,072	124	4701	4257	1463	1433
A318	0,0643	114	4574	3740	1280	1219
F50	0,1157	50	818	1243	1164	1150

Elaboración propia con base en Course Packet, Airline Operations, Purdue University

L: conjunto original de vuelos programables: El conjunto L del caso de aplicación tiene 432 vuelos, cada uno caracterizado por su número de vuelo, origen, destino, hora de salida y hora de llegada. Éstos se muestran en el Anexo B. Las horas de salida y llegada de cada vuelo fueron convertidas a formato decimal para poder utilizarlas en GAMS.

L^F : conjunto de vuelos obligatorios: De los 432 vuelos, se asumieron como obligatorios 210.

L^O : conjunto de vuelos opcionales (susceptibles de ser eliminados): Los 222 vuelos restantes son los que se clasificaron como opcionales.

El conjunto L^R para este caso de aplicación tiene 43 vuelos (sombreados en el Anexo B)

Parámetros para crear copias de cada vuelo

En este caso de aplicación se supondrá $a = 2$ (*número de copias*) y $b = 1$ (*longitud de la ventana de tiempo*). Esto implica que cada vuelo tendrá 3 versiones: la original, una copia 1 (saliendo y llegando media hora antes de la hora original) y una copia 2 (saliendo y llegando media hora después de la hora original). En el caso de aplicación se mantiene la misma notación expuesta en el caso básico de estudio (l^- para los vuelos que salen antes de la hora original, l_0 para los vuelos originales y l^+ para los vuelos que salen después de la hora original).

Conjunto ampliado de vuelos

V es el conjunto ampliado de vuelos. Como hay 432 vuelos originales, V tiene $1296 = 432 \times 3$ vuelos.

Otros conjuntos

N : conjunto de nodos de la red en el tiempo y en el espacio de cada tipo de avión.

De la misma manera que en el caso básico de estudio, para cada vuelo en el conjunto V se crean los nodos de llegada y salida. Como el conjunto V tiene 1296 vuelos, el conjunto N tendrá $2592 = 1296 \times 2$ nodos. Cada aeropuerto tiene un subconjunto de nodos, ordenados cronológicamente.

$U(N)$: conjunto de últimos nodos del día en cada aeropuerto: El conjunto U en el caso de aplicación cuenta con 42 elementos, uno para cada estación representando el último evento que ocurre en el día en ella (despegue o aterrizaje).

$P(N)$: conjunto de primeros nodos del día en cada aeropuerto: El conjunto P , al igual que U , tiene 42 elementos, uno para cada estación, a la primera hora del día que ocurre un evento en ella (un despegue o un aterrizaje).

Parámetros

C_{fv} : costo de asignar el tipo de avión f al vuelo v

Como se mencionó en las suposiciones, el costo de operar un vuelo entre dos ciudades se calculará de forma proporcional a la distancia entre ellas según las ecuaciones (4.1), (4.2) y (4.3). En la Tabla 4-1 se muestran el CASM y el número de sillas de cada uno de los aviones de la flota. Las distancias entre cada par de ciudades se muestran en el Anexo B. Para calcular el costo de oportunidad, se asumió un $RASM = 0.15$ dólares y una demanda d_l para cada vuelo l . Esta demanda también se muestra en el Anexo B.

A_f : número de aviones disponibles del tipo f .

El número de aviones de cada tipo según se muestra en la Tabla 4-2.

Tabla 4-2: Número de aviones disponibles por tipo en el caso de aplicación

Tipo de avión	Número de aviones disponibles
A330	7
B767	5
A320	28
A319	16
A318	10
F50	10

Fuente: Informe de gestión Avianca 2011.

R_l : pérdida de rentabilidad si el vuelo l se excluye del itinerario final.

Como se mencionó en las suposiciones, la rentabilidad asociada a cada vuelo se calculará como un porcentaje sobre el costo, según la ecuación (4.4). Al igual que en el caso básico de estudio, la pérdida de rentabilidad se calcula para los vuelos originales y no para los vuelos copia.

Los parámetros $Conect_{vn}$, K_{lv} , HS_v , HL_v , $Salida_l$, $Llegada_l$ y ρ se calculan de la misma forma que en el caso básico de estudio.

La restricción de factibilidad de la operación (que asegura que cada vuelo se cubra con un tipo de avión *apropiado*, se modelizará en este caso de aplicación introduciendo un

parámetro w_{fv} que tomará el valor de 1 si el avión de tipo f no está habilitado para cubrir el vuelo v y 0 en caso contrario. Así, la restricción se muestra en la ecuación (4.7).

$$X_{fv} = 0 \text{ cuando } w_{fv} = 1 \quad (4.7)$$

El parámetro w_{fv} tomará el valor de 1 si se cumple cualquiera de las siguientes condiciones:

- La distancia entre las ciudades servidas por el vuelo v es mayor que el alcance máximo del tipo de avión f .
- La longitud de pista de las ciudades servidas por el vuelo v es menor que la longitud mínima requerida para despegar o aterrizar por el tipo de avión f .
- El avión de tipo f está limitado para vuelos nacionales y el vuelo v es internacional. (Esta condición aplica únicamente para los F50 y A318).

Para calcular el valor del parámetro w_{fv} , se usaron las distancias entre ciudades mostradas en el Anexo B, y los rangos y longitudes de pistas requeridas por cada tipo de avión mostradas en la Tabla 4-1.

4.5 Programación del modelo de optimización sin clusterización

El modelo de optimización del caso aplicación sin involucrar la clusterización de datos se programó en GAMS 24.7.3. El archivo .gms se puede descargar en el siguiente [vínculo](#). Debido a su tamaño (en términos de la cantidad de variables y ecuaciones), el modelo se corrió en la plataforma NEOS Server, usando un servidor del Departamento de Ciencias de la Computación de la Universidad de Wisconsin-Madison.

4.6 Resultados del caso de aplicación sin clusterización

De manera similar al caso básico de estudio, para este modelo interesa saber cuáles vuelos fueron incluidos en el itinerario, qué copia de los vuelos resultó elegida, a qué tipo de avión fue asignado cada vuelo, cuántas conexiones exitosas se lograron establecer y cuál es el valor óptimo de la función objetivo.

4.6.1 Resultados de las variables de decisión

De los 432 vuelos, se asignaron 405. La distribución de vuelos asignados a cada tipo de avión, así como la comparación de la cantidad de aviones usados vs los disponibles, se muestra en la Tabla 4-3.

Tabla 4-3: Resultados del caso de aplicación sin clusterizar los datos

Tipo de avión	Cantidad de vuelos asignados	Número de aviones usados	Número de aviones disponible
A330	18	7	7
B767	12	5	5
A320	230	28	28
A319	53	16	16
A318	42	10	10
F50	50	10	10

Elaboración propia

Como se puede ver, se usó la totalidad de la flota disponible. En el Anexo C se puede encontrar la asignación detallada de cada vuelo. Esta selección de vuelos se opera con un costo mínimo total de \$5'653.081 dólares, desagregado en la Tabla 4-4. Este itinerario genera un total de 16.878 conexiones exitosas entre vuelos.

Tabla 4-4: Discriminación de costos de la solución óptima del caso de aplicación sin clusterizar los datos

Tipo de costo	Valor
Costo operativo	\$2.514.272
Pérdida de rentabilidad por exclusión de vuelos del itinerario	\$79.934
Penalización por conexiones infactibles	\$3.058.875
COSTO TOTAL	\$5.653.081

Elaboración propia

Al comparar las horas de salida y los tipos de avión asignados a los vuelos incluidos en la solución del modelo propuesto respecto al itinerario original de Avianca, se observó una coincidencia del 13% en las horas de salida y del 36% en la asignación del tipo de avión (tomando como base los 405 vuelos que fueron elegidos por el modelo). Esto indica que el modelo propuesto tomó partido de la flexibilidad dada en cuanto a las horas de salida y asignó aquellas que permitieran maximizar la cantidad de conexiones exitosas entre

vuelos. La coincidencia en la asignación del tipo de avión puede explicarse por los parámetros usados para calcular los costos de asignación: hay algunos vuelos que, para asemejarse a la elección original de tipo de avión, requerirían un cálculo de parámetros basado en datos reales de la aerolínea.

4.6.2 Tamaño del modelo sin clusterización

En la Tabla 4-5 se resumen algunas características del modelo, en términos de la cantidad de variables y de ecuaciones, y el tiempo de procesamiento y solución del modelo, usando los datos del caso de aplicación y resolviendo el problema usando la plataforma Neos Server.

Tabla 4-5: Estadísticas del modelo usando los datos del caso básico de aplicación

Número de variables de decisión	64.682
Número de ecuaciones (Función objetivo y restricciones)	98.693
Número de iteraciones requeridas para llegar a la solución óptima	3'633.494
Tiempo de solución requerido por el servidor disponible	1,28 horas

Elaboración propia

De la Tabla 4-5, y como se veía desde el caso básico de estudio, se destaca la cantidad de variables de decisión y de ecuaciones necesarias para decidir sobre los 432 vuelos: 64.682 variables de decisión (binarias y enteras) y 98.693 ecuaciones. Esta cantidad de variables, así como el tiempo de solución requerido, son una muestra del carácter NP-hard de este problema de optimización.

4.7 Formulación algebraica del caso de aplicación involucrando clusterización de datos

La cantidad de variables y de ecuaciones requeridas para formular el problema de optimización del caso de aplicación, y el tiempo requerido por el solver para entregar la solución óptima son una evidencia de la necesidad de usar una estrategia que permita reducir el tamaño del problema. En esta sección se implementa el algoritmo *k-Means* de clusterización de datos explicado en los Capítulos 2 y 3, para formular el modelo de optimización con el conjunto F clusterizado.

4.7.1 Procedimiento de clusterización del conjunto F de tipos de avión

Los tipos de aviones con los que se trabajó el caso de aplicación sin clusterizar los datos son los siguientes:

- Airbus A330 (A330)
- Boeing 767 (B767)
- Airbus A320 (A320)
- Airbus A319 (A319)
- Airbus A318 (A318)
- Fokker 50 (F50)

Como se mostró en la Tabla 4-1, se consideraron seis parámetros técnicos para caracterizar cada tipo de avión: CASM (dólares), número de sillas, costo operativo por hora, rango de alcance (millas), la longitud de pista requerida para despegue (m) y longitud de pista requerida para aterrizaje (m).

Entradas del algoritmo de clusterización de datos

- Tipos de aviones originales (6)
- Características técnicas de cada tipo de avión (6)
- Número de clústers (2)

Salidas del algoritmo de clusterización de datos

- Asignación de cada tipo de avión a uno de los dos clústers
- Centroides de los clústers (expresados en la escala original de los datos)

De acuerdo con el pseudocódigo presentado en el Capítulo 3, se programó el algoritmo de clusterización en R. El código se encuentra en el Anexo D.

Después de correr el algoritmo usando los datos de la Tabla 4-1 se obtuvieron dos clústers: uno que contiene los tipos A330 y B767, y otro que contiene los tipos A320, A319, A318 y F50. En la Tabla 4-6 se aprecian los centroides de cada clúster, ya expresados en sus unidades originales.

Tabla 4-6: Centroides de los clústers de tipos de avión resultantes

Clúster	Tipos de aviones que lo integran	CASM (dólares)	Número de sillas	Costo operativo por hora (dólares)	Alcance (millas)	Longitud de pista requerida para despegue (m)	Longitud de pista requerida para aterrizaje (m)
1	A330 B767	0.05385	279	8730	7960	2420	1637
2	A320 A319 A318 F50	0.078675	109.5	3621.5	3187.25	1401.75	1316.25

Elaboración propia usando R

Una característica física evidente en los tipos de aviones pertenecientes a cada clúster es que los aviones del clúster 1 son aviones de doble pasillo y los aviones del clúster 2 son aviones de pasillo sencillo. En este sentido, los nombres asignados a cada clúster a partir de este punto serán **pDoble** y **pSencillo**. Así, el conjunto F inicialmente representado en la ecuación (4.6) se cambia por un conjunto con únicamente dos elementos: **pDoble** y **pSencillo**, como se ve en la ecuación (4.8).

$$F = \{pDoble, pSencillo\} \quad (4.8)$$

4.7.2 Cálculo de parámetros que dependen del tipo de avión

Como ya no se tienen los tipos de avión de manera desagregada sino clusterizados, los parámetros técnicos que se usarán para calcular los parámetros del modelo corresponderán a los obtenidos en los centroides de los clústers que se mostraron en la Tabla 4-6. Por ejemplo, el CASM de los aviones de pasillo doble que se usará para calcular los costos de asignación será de \$0.05385. El número de aviones disponibles en cada clúster corresponderá a la suma de aviones que integran el clúster. Esta agregación se muestra en la Tabla 4-7.

Tabla 4-7: Número de aviones disponibles por clúster

Tipo de avión	Número de aviones disponibles
pDoble	12
pSencillo	64

Elaboración propia

Para calcular el valor del parámetro w_{fv} , se usaron de nuevo las distancias entre ciudades mostradas en el Anexo B, y los rangos y longitudes de pistas calculadas para cada clúster mostradas en la Tabla 4-6.

4.8 Programación del modelo de optimización con clusterización de los tipos de aviones

El modelo de optimización del caso aplicación involucrando la clusterización de los tipos de aviones se programó también en GAMS 24.7.3 (modificando el caso sin clusterizar). El archivo .gms se puede descargar en el siguiente [vínculo](#). A pesar de ser un modelo de menor tamaño (por involucrar un conjunto de menor tamaño), también tuvo que ser corrido en la plataforma NEOS Server, usando de nuevo un servidor del Departamento de Ciencias de la Computación de la Universidad de Wisconsin-Madison.

4.9 Resultados del caso de aplicación con clusterización

De forma análoga al caso básico de estudio y al modelo sin involucrar clusterización, los resultados que interesan son los vuelos incluidos en el itinerario (y la copia elegida), el clúster de aviones asignado a cada vuelo, el número de conexiones exitosas que se lograron establecer y cuál es el valor óptimo de la función objetivo. Respecto al valor de la función objetivo, debe tenerse en cuenta que no es comparable directamente con el obtenido en el caso sin clusterización, debido a que los costos unitarios calculados para este caso con clusterización corresponden al promedio de los costos unitarios de todos los tipos de aviones pertenecientes al clúster, y por ello no consisten en un costo real asignable al vuelo, sino más bien un costo representativo usado como criterio de elección entre los clústers. Más adelante se muestra el cálculo detallado del costo total de la operación involucrando clusterización.

4.9.1 Valores de las variables de decisión

De los 432 vuelos, se asignaron de nuevo 405, aunque distintos a los elegidos en el modelo sin clusterización. La distribución de vuelos asignados a cada clúster de aviones, así como la comparación de la cantidad de aviones usados vs los disponibles, se muestra en la Tabla 4-8.

Tabla 4-8: Resultados del caso de aplicación al clusterizar los datos

Tipo de avión	Cantidad de vuelos asignados	Número de aviones usados	Número de aviones disponible
pDoble	38	12	12
pSencillo	367	64	64

Elaboración propia

Como se puede ver, de nuevo se utilizó la totalidad de la flota disponible. En el Anexo C se puede encontrar la asignación detallada de cada vuelo a cada clúster de tipos de aviones. Este itinerario genera un total de 17.304 conexiones exitosas entre vuelos. Respecto al modelo sin involucrar clusterización, este modelo asegura 426 conexiones exitosas más, debido a que, respecto al modelo sin clusterización hay cambios en la hora de salida de 91 vuelos, así: 83 vuelos salen a horas diferentes, se incluyen 4 vuelos que no estaban en el modelo sin clusterización y se excluyen 4 vuelos que estaban en el modelo sin clusterización.

Al comparar las horas de salida y los tipos de avión asignados a los vuelos incluidos en la solución del modelo propuesto con clusterización respecto al itinerario original, se observó una coincidencia del 9% en las horas de salida y del 95% en la asignación del tipo de avión. La razón del incremento del porcentaje de coincidencia para los tipos de avión es que en esta ocasión se clasificó como coincidencia que el tipo de avión asignado en el itinerario original perteneciera al clúster elegido por el modelo.

4.9.2 Tamaño del modelo involucrando clusterización

En la Tabla 4-9 se resumen algunas características del modelo, en términos de la cantidad de variables y de ecuaciones, y el tiempo de procesamiento y solución del modelo, usando los datos del caso de aplicación y resolviendo el problema usando la plataforma Neos Server.

Tabla 4-9: Comparación entre las estadísticas del modelo usando los datos del caso básico de aplicación con y sin clusterización de datos

Tipo de problema	Sin clusterización de datos (SC)	Con clusterización de datos (CC)	Razón SC/CC
Número de variables de decisión	64.682	49.130	1,3
Número de ecuaciones (Función objetivo y restricciones)	98.693	88.321	1,1
Número de iteraciones requeridas para llegar a la solución óptima	3'633.494	154.319	24
Tiempo de solución requerido por el servidor disponible	1,28 h	210 s	22

Elaboración propia

La comparación presentada en la Tabla 4-9 revela el carácter NP-hard del problema planteado. El problema sin clusterización de datos es 1,3 veces más grande que el que sí la involucra en términos de las variables de decisión y 1,1 veces más grande en términos de la cantidad de ecuaciones. Sin embargo, la cantidad de iteraciones necesarias para llegar a la solución óptima es 24 veces más grande, y el tiempo de solución es 22 veces más grande.

4.10 Implementación del cambio de escala para asignar a cada vuelo elegido un tipo de avión

Como se explicó, hasta este punto se tiene una asignación preliminar de cada vuelo incluido en el itinerario a un clúster de aviones. Para poder entregar una respuesta en la escala original, se corre el modelo reducido presentado en el Capítulo 3 para cada uno de los dos clústers. Estos modelos fueron programados en GAMS. Los archivos .gms pueden consultarse [aquí](#) para el clúster pDoble y [aquí](#) para el clúster pSencillo. Así, se logra obtener la solución del problema en su escala original, puesto que con la solución del modelo reducido para cada clúster se obtiene un tipo de avión específico asignado de manera factible a cada uno de los 405 vuelos incluidos en el itinerario. En la Tabla 4-10 se muestran las estadísticas de los modelos.

Tabla 4-10: Estadísticas de los modelos reducidos planteados para desclusterizar la solución agregada

Clúster	pDoble	pSencillo	Suma
Número de variables de decisión	229	4.405	4.634
Número de ecuaciones (Función objetivo y restricciones)	193	3.308	3501
Número de iteraciones requeridas para llegar a la solución óptima	20	427	447
Tiempo de solución requerido por el servidor disponible	0,01 s	0,178 s	0,188 s

Elaboración propia

La distribución de vuelos asignados a cada clúster de aviones, así como la comparación de la cantidad de aviones usados vs los disponibles, se muestra en la Tabla 4-11.

Tabla 4-11: Resultados del caso de aplicación al desclusterizar los tipos de avión según los resultados de los modelos reducidos

Tipo de avión	Cantidad de vuelos asignados	Número de aviones usados	Número de aviones disponible
A330 (pDoble)	18	7	7
B767 (pDoble)	20	5	5
A320 (pSencillo)	216	28	28
A319 (pSencillo)	68	16	16
A318 (pSencillo)	35	10	10
F50 (pSencillo)	48	10	10

Elaboración propia

Como se puede ver, de nuevo se utilizó la totalidad de la flota disponible. En el Anexo C se puede encontrar la asignación detallada de cada vuelo a cada clúster de tipos de aviones.

4.11 Comparación general de resultados

En las secciones anteriores, en la medida en que se fueron obteniendo soluciones de los modelos, se hicieron algunas comparaciones. Aquí se retoman los resultados obtenidos con cada enfoque y se comparan los resultados de los modelos con y sin clusterización en los siguientes aspectos:

- Costos operativos, pérdida de rentabilidad y penalización por conexiones infactibles.

Tabla 4-12: Costos operativos, pérdida de rentabilidad y penalización por conexiones infactibles en los modelos con y sin clusterización

Tipo de costo	Itinerario base ¹	Modelo propuesto sin clusterización	Modelo propuesto con clusterización
Costo operativo	\$2.865.063	\$2.514.272	\$2.531.978 ²
Pérdida de rentabilidad por exclusión de vuelos del itinerario	\$0	\$79.934	\$84.066
Penalización por conexiones infactibles	\$3.010.875	\$3.058.875	\$3.005.625
COSTO TOTAL (Dólares)	\$5.875.938	\$5.653.081	\$5.621.669

Elaboración propia

En la Tabla 4-12 se muestra uno de los aportes fundamentales del modelo y de la metodología de clusterización propuesta: **la selección de aviones y vuelos no solo disminuyó la penalización por incluir conexiones infactibles (respecto al itinerario base y al modelo sin clusterización), sino que disminuyó el costo operativo (asociado a la asignación final de tipos de aviones a vuelos). Esto indica que la flexibilidad dada al permitir la selección de la hora de salida del vuelo y al clusterizar los tipos de avión permitió obtener una asignación más económica desde la perspectiva financiera, y por supuesto factible desde la perspectiva operacional.** Un ejercicio que podría plantearse, en el caso de querer utilizar el modelo propuesto para una aerolínea real, sería comparar el itinerario base con el itinerario propuesto por los modelos con y sin clusterización, utilizando valores de parámetros que reflejen de una manera más cercana la realidad de la operación de la aerolínea.

¹ Suponiendo la inclusión de los 432 vuelos originalmente planteados y la asignación al tipo de avión reportado en el itinerario.

² Obtenido de sumar los costos de operación en los dos modelos reducidos.

- Número de variables de decisión y ecuaciones, cantidad de iteraciones requeridas para llegar a la solución óptima y tiempo total de solución.

Tabla 4-13: Comparación de los tamaños de los modelos con y sin clusterización

Modelo	Modelo propuesto sin clusterización (SC)	Modelo propuesto con clusterización (CC) ³	Razón SC/CC
Número de variables de decisión	64.682	53.535	1,2
Número de ecuaciones (Función objetivo y restricciones)	98.693	91.629	1,1
Número de iteraciones requeridas para llegar a la solución óptima	3.633.494	154.746	23
Tiempo de solución requerido por el servidor disponible	1,27 horas	210,178 s	22

Elaboración propia

De la Tabla 4-13 se puede ver que incluir los dos modelos reducidos no aumenta significativamente el tamaño del modelo con clusterización (según se mostró en la Tabla 4-9) y que su tiempo de solución total es mucho menor que el del modelo sin clusterización. **Así, tanto por la calidad de la solución, como por su tiempo de obtención, puede verse que, por lo menos para el conjunto de datos analizado, la clusterización de datos es una metodología útil para realizar la solución integrada de las dos etapas de planeación de la operación.**

- Similitud entre las horas de salida y los tipos de avión asignados respecto al itinerario original de Avianca.

Tabla 4-14: Similitud entre las horas de salida y los tipos de avión asignados respecto al itinerario original de Avianca

Modelo	Modelo propuesto sin clusterización	Modelo propuesto con clusterización
% de similitud respecto a las horas de salida de los vuelos en el itinerario de Avianca	13%	9%
% de similitud respecto al tipo de avión asignado en el itinerario de Avianca	36%	35%

Elaboración propia

³ Calculados sumando el modelo global con clusterización de tipos de avión con los dos modelos reducidos.

Como se había mencionado anteriormente, los porcentajes de similitud reportados en la Tabla 4-14 (con base en los 405 vuelos elegidos por los modelos) dan cuenta de que el modelo propuesto en esta tesis efectivamente aprovecha la flexibilidad dada para elegir la mejor hora de salida de cada vuelo y privilegiar aquellas que favorecen las conexiones exitosas. Respecto a la selección del tipo de avión y como también se mencionó antes, los porcentajes obtenidos dan cuenta de una necesidad de calibrar los parámetros relacionados con los costos operativos (CASM y demanda de pasajeros de cada vuelo) si se desea implementar este modelo en un contexto real.

- Similitud entre las horas de salida y los tipos de avión asignados en el modelo con clusterización vs el modelo sin clusterización:

Tabla 4-15: Similitud entre las horas de salida y los tipos de avión asignados en el modelo con clusterización vs el modelo sin clusterización

% de similitud respecto a las horas de salida de los vuelos	84%
% de similitud respecto al tipo de avión asignado	81%

Elaboración propia

En la Tabla 4-15 se aprecia que el 84% de los vuelos conservaron la misma hora de salida en los modelos con y sin clusterización, y que el 81% de los vuelos conservaron el mismo tipo de avión asignado. La diferencia puede explicarse a partir del hecho de que en el modelo clusterizado, al contar con un grupo de aviones (en lugar de tipos individuales), se contó con más flexibilidad para privilegiar las conexiones exitosas sobre la selección del tipo de avión.

- Número de conexiones exitosas establecidas.

Tabla 4-16: Número de conexiones exitosas establecidas por los modelos con y sin clusterización

	Itinerario base ⁴	Modelo propuesto sin clusterización	Modelo propuesto con clusterización
Número de conexiones exitosas establecidas	17.262	16.878	17.304

Elaboración propia

Los resultados de la Tabla 4-16 son consistentes con los de la Tabla 4-15, en los que se ve que una agrupación de aviones por clúster permitió al modelo buscar favorecer las

⁴ Suponiendo la inclusión de los 432 vuelos originalmente planteados.

conexiones exitosas, dado que la asignación del tipo específico de avión a cada vuelo no se realizó sino hasta la implementación del modelo reducido. Además, el hecho de que haya más conexiones exitosas en el modelo con clusterización respecto al itinerario base (incluso excluyendo vuelos), de nuevo da cuenta del aprovechamiento de la flexibilidad dada al modelo para elegir la hora de despegue de cada vuelo.

- Número de vuelos asignados a cada tipo de avión.

Tabla 4-17: Número de vuelos asignados a cada tipo de avión en los modelos con y sin clusterización

Tipo de avión	Modelo propuesto sin clusterización	Modelo propuesto con clusterización
A330	18	18
B767	12	20
A320	230	216
A319	53	68
A318	42	35
F50	50	48

Elaboración propia

A partir de los resultados de la Tabla 4-17, vale la pena resaltar que la selección del tipo de avión estaba condicionada por tres elementos: el costo operativo, la disponibilidad de aviones y la factibilidad de la operación. En el modelo sin clusterización cada elemento se trabajó de manera explícita y directa: el costo operativo se calculó con el CASM real de cada avión, se usó la cantidad disponible de aviones de cada tipo, y se aplicaron las condiciones de infactibilidad mencionadas (longitud de pista, rango de alcance y aviones limitados a vuelos nacionales). En el modelo con clusterización, sin embargo, estas tres condiciones se relajaron para facilitar la flexibilidad del modelo: se usó un CASM promedio para cada clúster, las disponibilidades de aviones se calcularon agrupando los aviones asignados a cada tipo de clúster, y las factibilidades de calcularon de manera genérica (considerando los alcances y longitudes de pista requeridas para despegue y aterrizaje promedio del clúster). Solo fue hasta la implementación de los modelos reducidos que se volvieron a calcular los costos de operación reales, y que se volvieron a aplicar las condiciones de factibilidad de la operación originales. Estos cambios se ven reflejados en la asignación final de aviones a vuelos, favoreciendo al A330 (en el clúster pDoble) y al A320 (en el clúster pSencillo).

5. Conclusiones y trabajos futuros

Sobre los modelos de optimización para el diseño de itinerarios y la asignación del tipo de avión

- El transporte urbano y el transporte aéreo de pasajeros guardan importantes diferencias que determinan la forma en que se planea y administra su operación: el transporte aéreo es de propiedad generalmente privada, y por ello muchas de las decisiones que se toman en los procesos de planeación están orientados a maximizar la rentabilidad y a disminuir los costos. En el transporte urbano, no necesariamente se deben optimizar criterios financieros sino que generalmente prima una maximización de la calidad del servicio prestado a los usuarios. A pesar de estas diferencias, en la academia puede observarse cómo para planear la operación de ambos modos de transporte, típicamente se ha hecho una división por etapas de esta planeación. Estas etapas son parecidas en cada uno de los dos modos de transporte y resuelven preguntas similares, por supuesto con la especificidad de cada modo.
- En el caso del diseño de itinerarios y de la asignación del tipo de avión, que son las primeras dos etapas de la planeación de la operación de una aerolínea, se encontraron modelos de optimización tanto para cada etapa por separado como para las dos etapas de manera integrada. Estos dos *extremos*: modelos secuenciales vs modelos integrados, responden a dos complejidades que se encuentran al planear la operación de sistemas de transporte: cuando se trabaja con modelos secuenciales, es posible que la solución global que se encuentre (como la suma de soluciones locales) no resulte óptima o factible. Debido a esto, ha surgido la tendencia de integrar etapas de la planeación en un solo modelo, con el propósito de obtener una sola solución que sí sea óptima (y por supuesto factible) para las etapas que se están considerando. Sin embargo, esta agrupación implica un aumento de la cantidad de variables de decisión, lo que, en el caso real de una

aerolínea (con miles de vuelos y cientos de aviones) puede imposibilitar la ejecución del modelo de optimización en un solver, debido a que se sobrepasa la capacidad de cómputo de éste. Dado este panorama, lo que se desea es obtener las mejores características de los ambos enfoques: la obtención rápida de soluciones que posibilitan los enfoques secuenciales y la obtención de una solución global para las etapas que se deseen integrar. Por ello se propuso en esta tesis un enfoque de cambio de resolución.

- Los modelos de diseño de itinerarios de la industria aeronáutica son un desafío para la investigación. En la literatura revisada se encontró que en su gran mayoría, la definición de itinerarios es considerada la elección de vuelos a programar de una lista de vuelos que pueden añadirse o eliminarse de un itinerario base con el que se cuenta previamente. Esto significa que en realidad el diseño de itinerarios implica una mejora respecto al itinerario base que se está trabajando, y que la decisión la frecuencia de vuelos entre cada par origen-destino se determina de forma indirecta al elegir los vuelos incluidos en el itinerario final, y que la hora de salida de cada vuelo puede elegirse entre un conjunto discreto de posibilidades. Hay varias causas que explican esta reducción de flexibilidad en el planteamiento de los modelos: se han encontrado dificultades en modelizar la forma como los pasajeros eligen volar y no se cuenta con datos benchmark que permitan comparar propuestas de modelos. En este contexto, desarrollar un modelo de diseño de itinerarios que considere la hora de salida de cada vuelo como una variable de decisión continua es una oportunidad de trabajo futuro. Esto permitiría ahorrarse la carga computacional que implica el uso de variables binarias.

Sobre parámetros de los modelos

- Los modelos de optimización para la planeación de la operación de sistemas de transporte dependen altamente de los valores de los parámetros. En el caso particular del modelo integrado para la definición de itinerarios y asignación de tipo de avión, los parámetros que más involucrados están en el proceso de decisión son los costos de operación de cada vuelo en cada tipo de avión, la pérdida de rentabilidad por dejar de incluir un vuelo del itinerario y la demanda de pasajeros (o de sillas) de cada vuelo. Estos parámetros, si bien pueden calcularse haciendo simplificaciones, dependen de factores muchas veces externos a la aerolínea o que no pueden conocerse con suficiente antelación (por ejemplo, el itinerario de la

competencia). Esto puede ser uno de los causantes de que los modelos típicos de la literatura no busquen hacer una definición desde cero, sino que se basen en itinerarios previos, como una forma de ganar incertidumbre sobre algunos parámetros.

Sobre la metodología de clusterización de datos

- La clusterización de datos es una herramienta de Inteligencia Artificial que se ha usado en distintos contextos, con el propósito de representar un conjunto de datos homogéneos entre sí por medio de un centroide. Esto disminuye la cantidad de observaciones que se necesitan para representar un conjunto de datos, así como las conexiones y relaciones entre ellos. Dentro de las distintas aplicaciones de clusterización se mencionaron la división de clientes por áreas geográficas para entrega de correo, la agrupación de tipos de vegetaciones para estudios de hábitat, la clasificación de píxeles para reconocimiento de patrones, el diseño de sistemas de distribución y la asignación de clientes a diferentes segmentos de mercadeo de acuerdo con sus características demográficas. En el desarrollo de esta tesis se encontró la oportunidad de vincular esta metodología de forma novedosa en el contexto de los modelos de optimización, para aprovechar la reducción de la cantidad de observaciones que produce la clusterización como una estrategia que permitiera cambiar la escala del modelo de optimización propuesto.

Sobre el modelo de optimización desarrollado en esta tesis

- El modelo desarrollado en esta tesis busca asignar de forma simultánea los vuelos que serán incluidos en el itinerario final de una aerolínea, y el tipo de avión asignado a dichos vuelos. Así, se resuelven de manera integrada dos de las etapas de la planeación de la operación de las aerolíneas, lo cual garantiza, por lo menos para esas dos etapas, que la solución propuesta cumple con el criterio de optimalidad. Sin embargo, el costo de hacer dicha integración, y como se mencionó a lo largo de todo el documento, es el aumento del tamaño del modelo, en términos de sus variables de decisión y ecuaciones y del tiempo de solución del modelo, tanto en tiempo como en cantidad de iteraciones. El caso básico de estudio, desarrollado para un conjunto de 36 vuelos, requirió el uso de 975 variables de decisión y 1.121 ecuaciones. El caso de aplicación, que consideró 432 vuelos, requirió (sin usar la metodología de clusterización) 64.382 variables de decisión y 98.693 ecuaciones. El tiempo de solución pasó de 0,014 segundos para realizar 881 iteraciones (para el caso básico de estudio) a 1,28 horas en las que se realizaron 3'633.494

iteraciones (en el caso de aplicación). Como puede verse, este aumento en el tiempo es no polinomial respecto al aumento del tamaño del problema, lo cual justifica que, dado el caso en el que una aerolínea real desee usar este modelo o que simplemente se desee usar un conjunto más grande de vuelos, se implementen técnicas que permitan disminuir el tiempo de solución.

- Uno de los parámetros más sensibles que se encontró en la formulación del modelo de esta tesis fue el escalar M usado para penalizar las conexiones infactibles entre cada par de vuelos analizados. El valor de M , en la literatura típica de formulación de problemas de optimización, se supone de varios órdenes de magnitud más grande que el resto de los valores del problema y no se le asigna un valor particular; esto se hace con el propósito de castigar drásticamente la inclusión en la solución óptima de cualquier variable penalizada por esta M . En el modelo propuesto, se incluye la M para castigar a los pares de vuelos con más conexiones infactibles que son incluidos en el itinerario final. Este castigo debe ser lo suficientemente severo para excluir del itinerario final los vuelos más inconvenientes (según su cantidad de conexiones infactibles), pero no de forma exagerada, porque si toma un valor demasiado alto, que un vuelo tenga o no conexiones infactibles con los demás vuelos será indiferente para el modelo y no habría manera de penalizar cada vuelo en función de sus conexiones infactibles, sino que simplemente se le excluiría del itinerario. De forma empírica se encontró que existen distintos rangos de M en los que el itinerario resultante es insensible al valor de este parámetro; sin embargo (y como es esperable), a mayor valor de M , más vuelos opcionales son excluidos del itinerario final.

Sobre la metodología de clusterización aplicada en esta tesis

- Con el propósito de disminuir el tiempo de solución del modelo de optimización para resolver de forma integrada las etapas de diseño de itinerarios y asignación de flota, se planteó la metodología de clusterización de datos. Usando algoritmos de Inteligencia Artificial y Aprendizaje de Máquina, se propuso agrupar los elementos de conjuntos de datos en k categorías o clústers. En el caso de aplicación desarrollado se utilizó el conjunto de tipos de avión. Este se dividió en 2 categorías o clústers, de manera que los tipos de avión pertenecientes a cada clúster fueran similares entre sí, y guardaran diferencias respecto a los tipos de aviones de los otros clústers. Las características técnicas que se usaron para diferenciar los tipos

de aviones fueron: el CASM, el número de sillas, el costo operativo por hora, rango de alcance, la longitud de pista requerida para despegue y la longitud de pista requerida para aterrizaje. En el algoritmo se usó la fórmula de la distancia con métrica 2 para calcular las distancias entre cada observación. Después de correr repetidamente el algoritmo, se eligió adoptar los clústers de **pDoble** (para los A330 y B767) y **pSencillo** (para los A320, A319, A318 y F50) debido a que fue el resultado que apareció con mayor frecuencia. Como se mencionó, la característica más predominante que permite identificar los clústers es la cantidad de sillas de cada tipo de avión, por lo que se ve que esta fue la variable más heterogénea al correr el algoritmo y que permitió separar los grupos de aviones.

- Cuando se aplica el algoritmo k-Means para agrupar datos, si se parte de centroides aleatorios, es natural que los clústers resultantes varíen entre múltiples corridas del algoritmo. En esta tesis, y aunque se menciona la cantidad de sillas como un patrón común en cada clúster, es necesario recordar que la idea que motivó el uso de esta herramienta de Inteligencia Artificial fue la de permitir al algoritmo agrupar los datos, en lugar de hacer una clasificación trivial que no calculara de manera cuantitativa la heterogeneidad entre cada tipo de avión.
- Debido a que los tipos de avión quedaron agrupados por clústers, sus características técnicas individuales pasaron a ser reemplazadas por las calculadas en el centroide de cada clúster. Esto hace que el óptimo calculado en el modelo aplicando clusterización sea representativo para los clústers y que luego de obtener los resultados finales del problema en su escala original, sea necesario calcular el costo total de la asignación resultante con los parámetros reales de cada tipo de avión.

Sobre el algoritmo de cambio de escala desarrollado en esta tesis para comparar los resultados de una escala vs otra

- La estrategia que se siguió para manejar los datos por clústers consistió en correr el modelo propuesto considerando un conjunto F de tipos de aviones con solo dos elementos: aviones de tipo **pDoble** y aviones de tipo **pSencillo**. Esto significa que la decisión que el modelo propuesto debía ayudar a tomar consistía en la selección de vuelos que harían parte del itinerario final, y de una asignación a cada uno de ellos de uno de los dos clústers de aviones. Esto significa que los resultados del modelo propuesto aplicando clusterización de datos quedaron expresados en una

escala *gruesa*, ya que no se asignó a cada vuelo un tipo *específico* de avión. Para poder llegar a una respuesta en las unidades originales con los que se planteó el problema, se optó por correr, para los vuelos asignados a cada uno de los dos clústers, una versión reducida del modelo propuesto, considerando únicamente las restricciones de balance, disponibilidad, cubrimiento obligatorio de vuelos y factibilidad, minimizando únicamente el costo operativo de asignación de aviones a vuelos. Así, se obtuvo una asignación de un tipo específico de avión a cada uno de los vuelos, garantizando la factibilidad de la asignación propuesta.

- La implementación de la metodología de clusterización para el caso de aplicación propuesto, evidenció dos ventajas respecto a correr un único modelo integrado: por un lado, el tiempo de solución utilizando la estrategia de clusterización (considerando el modelo general con el conjunto F clusterizado y luego los dos modelos reducidos para cada clúster) fue 22 veces menor que el del modelo sin clusterización. Por otro lado, la calidad de la solución al aplicar la metodología de clusterización fue superior a la de la solución sin aplicar clusterización: el costo total fue menor (pasó de \$5.653.081 a \$5.621.669 dólares), y se garantizó una mayor cantidad de conexiones exitosas (pasó de 16.878 a 17.304). Una de las razones que permite explicar esta mejora en la calidad de la solución tiene que ver con la flexibilidad que se le dio al modelo al agrupar los tipos de avión por clúster: esto permitió que el modelo contara con más opciones para garantizar conexiones exitosas entre vuelos. Así, puede considerarse que en esta tesis se hizo un aporte a la pregunta abierta descrita en la sección 1.2.6, en la que se indagaba sobre la interacción entre los modelos de optimización y herramientas matemáticas de manejo de datos.

Sobre la aplicabilidad del modelo y la metodología propuesta a problemas de otros contextos

- Aunque en el caso de aplicación desarrollado se decidió clusterizar el únicamente el conjunto de tipos de avión, esta propuesta puede extenderse a clusterizar otros conjuntos de datos. Por ejemplo, se podría agrupar el conjunto de aeropuertos de la red, considerando como criterios de clusterización, por ejemplo, la distancia al hub principal de la aerolínea, la distancia al centro demográfico del país, o cantidades como el volumen de pasajeros que maneja o de operaciones que administra. Esta propuesta se puede abordar en futuros trabajos. Al igual que en

esta tesis, para cada clusterización de datos debe pensarse en una “desclusterización” que al final del proceso entregue las respuestas del problema en sus unidades originales.

- Si bien en esta tesis se trabajó con base en los modelos de optimización para la planeación de la operación del transporte aéreo de pasajeros, la metodología de clusterización aplicada puede trasladarse por ejemplo a los modelos de optimización para la planeación del transporte urbano: debido a que en esa área la planeación también se ha hecho de manera secuencial, y se cuenta con modelos integrados para dos o más etapas, tiene sentido que con una agrupación de datos se disminuya la complejidad computacional de dichos modelos.

Sobre la relación entre el diseño de itinerarios y la complejidad de la red resultante

- Una oportunidad de trabajos futuros está en la vinculación de la teoría de complejidad de redes a la toma de decisiones en la definición del itinerario en una aerolínea: medidas en principio tan abstractas como las de centralidad, tamaño de paso y clusterización (calculadas para una red genérica) tienen vital importancia en la constitución de la red de rutas de una aerolínea y deberían por tanto ser incluidas como criterios que orienten la toma de decisiones en este contexto. Específicamente, es de especial interés poder incluir en los modelos de diseño de itinerarios, ecuaciones que permitan calcular las medidas de complejidad de la red resultante en función de los vuelos que incluyan en el itinerario final. Así, la decisión final satisfaría criterios de desempeño financieros, operativos, y también cumpliría con medidas deseadas de complejidad.

6. Bibliografía

- Abdelghany, A., Abdelghany, K., & Azadian, F. (2017). Airline flight schedule planning under competition. *Computers and Operations Research*, 87, 20–39. <http://doi.org/10.1016/j.cor.2017.05.013>
- Akartunal, K., Boland, N., Evans, I., Wallace, M., & Waterer, H. (2013). Airline planning benchmark problems-Part I:: Characterising networks and demand using limited data. *Computers and Operations Research*, 40(3), 775–792. <http://doi.org/10.1016/j.cor.2012.02.012>
- Bard, J. F., & Jarrah, A. I. (2009). Large-scale constrained clustering for rationalizing pickup and delivery operations. *Transportation Research Part B: Methodological*, 43(5), 542–561. <http://doi.org/10.1016/j.trb.2008.10.003>
- Barnhart, C., & Cohn, A. (2004). Airline Schedule Planning: Accomplishments and Opportunities. *Manufacturing & Service Operations Management*, 6(1), 3–22. <http://doi.org/10.1287/msom.1030.0018>
- Bazargan, M. (2010). *Airline Operations and Scheduling* (2nd ed.). Ashgate.
- Bélanger, N., Desaulniers, G., Soumis, F., Desrosiers, J., & Lavigne, J. (2006). Weekly airline fleet assignment with homogeneity. *Transportation Research Part B: Methodological*, 40(4), 306–318. <http://doi.org/10.1016/j.trb.2005.03.004>
- Belobaba, P., Odoni, A., & Barnhart, C. (2009). *The Global Airline Industry*. John Wiley & Sons. Retrieved from <http://books.google.com/books?hl=en&lr=&id=BRtDI0CJpQIC&pgis=1>
- Bešinović, N., Goverde, R. M. P., Quaglietta, E., & Roberti, R. (2016). An integrated micro-macro approach to robust railway timetabling. *Transportation Research Part B: Methodological*, 87, 14–32. <http://doi.org/10.1016/j.trb.2016.02.004>
- Brons, M., Pels, E., Nijkamp, P., & Rietveld, P. (2002). Price elasticities of demand for passenger air travel: a meta-analysis. *Journal of Air Transport Management*, 8(3), 165–175. [http://doi.org/https://doi.org/10.1016/S0969-6997\(01\)00050-3](http://doi.org/https://doi.org/10.1016/S0969-6997(01)00050-3)
- Cadarso, L., & de Celis, R. (2017). Integrated airline planning: Robust update of scheduling and fleet balancing under demand uncertainty. *Transportation Research Part C: Emerging Technologies*, 81, 227–245. <http://doi.org/10.1016/j.trc.2017.06.003>
- Cadarso, L., & Marín, Á. (2012). Robust passenger oriented timetable and fleet assignment integration in airline planning. *Journal of Air Transport Management*, 26, 44–49.

<http://doi.org/10.1016/j.jairtraman.2012.10.004>

- Ceder, A. (2007). *Public Transit Planning and Operation: Theory, Modeling and Practice* (Vol. 1). Burlington: Elsevier. <http://doi.org/10.1017/CBO9781107415324.004>
- Dong, Z., Chuhang, Y., & Lau, H. Y. K. H. (2016). An integrated flight scheduling and fleet assignment method based on a discrete choice model. *Computers and Industrial Engineering*, 98, 195–210. <http://doi.org/10.1016/j.cie.2016.05.040>
- Gu, Z., Johnson, E. L., Nemhauser, G. L., & Yinhu, W. (1994). Some properties of the fleet assignment problem. *Operations Research Letters*, 15(2), 59–71. [http://doi.org/https://doi.org/10.1016/0167-6377\(94\)90001-9](http://doi.org/https://doi.org/10.1016/0167-6377(94)90001-9)
- Guihaire, V., & Hao, J.-K. (2008). Transit network design and scheduling: A global review. *Transportation Research Part A: Policy and Practice*, 42(10), 1251–1273. <http://doi.org/10.1016/j.tra.2008.03.011>
- Harrington, P. (2015). *Machine Learning in Action. ... for Engineering and* Manning Publications. <http://doi.org/10.1007/978-0-387-77242-4>
- Härtel, P., Kristiansen, M., & Korpås, M. (2017). Assessing the impact of sampling and clustering techniques on offshore grid expansion planning. *Energy Procedia*, 137, 152–161. <http://doi.org/10.1016/j.egypro.2017.10.342>
- Ibarra-Rojas, O. J., Giesen, R., & Rios-Solis, Y. A. (2014). An integrated approach for timetabling and vehicle scheduling problems to analyze the trade-off between level of service and operating costs of transit networks. *Transportation Research Part B: Methodological*, 70, 35–46. <http://doi.org/10.1016/j.trb.2014.08.010>
- Jaramillo, P., & Lotero, L. (2010). *Modelos de optimización como apoyo a la planificación de la operación del transporte urbano público colectivo*.
- Kargari, M., & Sepehri, M. M. (2012). Stores clustering using a data mining approach for distributing automotive spare-parts to reduce transportation costs. *Expert Systems with Applications*, 39(5), 4740–4748. <http://doi.org/10.1016/j.eswa.2011.09.121>
- Lee, R. J., & Sener, I. N. (2016). Transportation planning and quality of life: Where do they intersect? *Transport Policy*, 48, 146–155. <http://doi.org/10.1016/j.tranpol.2016.03.004>
- Lohatepanont, M. (2002). Airline fleet assignment and schedule design: integrated models and algorithms, 38(1995), 195. <http://doi.org/10.1287/trsc.1030.0026>
- Marín, Á., & Jaramillo, P. (2008). Urban rapid transit network capacity expansion. *European Journal of Operational Research*, 191(1), 45–60. <http://doi.org/https://doi.org/10.1016/j.ejor.2007.08.010>
- Mesbah, M., Sarvi, M., Ouyeyi, I., & Currie, G. (2011). Optimization of transit priority in the transportation network using a decomposition methodology. *Transportation Research Part C: Emerging Technologies*, 19(2), 363–373. <http://doi.org/10.1016/j.trc.2010.05.020>
- Migdalas, A. (1995). Bilevel programming in traffic planning: Models, methods and challenge. *Journal of Global Optimization*, 7(4), 381–405.

- <http://doi.org/10.1007/BF01099649>
- Milone, D. H., Stegmayer, G., Kamenetzky, L., López, M., & Carrari, F. (2013). Clustering biological data with SOMs: On topology preservation in non-linear dimensional reduction. *Expert Systems with Applications*, 40(9), 3841–3845. <http://doi.org/10.1016/j.eswa.2012.12.074>
- Namratha, M., & Prajwala, T. R. (2012). A Comprehensive Overview of Clustering Algorithms in Pattern Recognition. *Journal of Computer Engineering*, 4(6), 23–30.
- Pai, V. (2010). On the factors that affect airline flight frequency and aircraft size. *Journal of Air Transport Management*, 16(4), 169–177. <http://doi.org/https://doi.org/10.1016/j.jairtraman.2009.08.001>
- Pita, J. P., Barnhart, C., & Antunes, A. P. (2012). Integrated Flight Scheduling and Fleet Assignment Under Airport Congestion. *Transportation Science*, 47(4), 477–492. <http://doi.org/10.1287/trsc.1120.0442>
- Poorman, J. P. (2005). A Holistic Transportation Planning Framework for Management and Operations.pdf. *ITE Journal*.
- Rodrigues, M. M., De Souza, C. C., & Moura, A. V. (2006). Vehicle and crew scheduling for urban bus lines. *European Journal of Operational Research*, 170(3), 844–862. <http://doi.org/10.1016/j.ejor.2004.06.035>
- Sahoo, A. K., Zuo, M. J., & Tiwari, M. K. (2012). A data clustering algorithm for stratified data partitioning in artificial neural network. *Expert Systems with Applications*, 39(8), 7004–7014. <http://doi.org/10.1016/j.eswa.2012.01.047>
- Schlechte, T., Borndörfer, R., Erol, B., Graffagnino, T., & Swarat, E. (2011). Micro-macro transformation of railway networks. *Journal of Rail Transport Planning and Management*, 1(1), 38–48. <http://doi.org/10.1016/j.jrtpm.2011.09.001>
- Sherali, H. D., Bae, K. H., & Haouari, M. (2010). Integrated airline schedule design and fleet assignment: Polyhedral analysis and benders' decomposition approach. *INFORMS Journal on Computing*, 22(4), 500–513. <http://doi.org/10.1287/ijoc.1090.0368>
- Sherali, H. D., Bish, E. K., & Zhu, X. (2006). Airline fleet assignment concepts, models, and algorithms. *European Journal of Operational Research*, 172(1), 1–30. <http://doi.org/10.1016/j.ejor.2005.01.056>
- Unternährer, J., Moret, S., Joost, S., & Maréchal, F. (2017). Spatial clustering for district heating integration in urban energy systems: Application to geothermal energy. *Applied Energy*, 190, 749–763. <http://doi.org/10.1016/j.apenergy.2016.12.136>
- Xie, L., Kliwer, N., & Suhl, L. (2012). Integrated Driver Rostering Problem in Public Bus Transit. *Procedia - Social and Behavioral Sciences*, 54, 656–665. <http://doi.org/10.1016/j.sbspro.2012.09.783>
- Xueyu, Z., & Jiaqi, Y. (2013). Research on the Bi-level Programming Model for Ticket Fare Pricing of Urban Rail Transit based on Particle Swarm Optimization Algorithm. *Procedia - Social and Behavioral Sciences*, 96(Cictp), 633–642. <http://doi.org/10.1016/j.sbspro.2013.08.074>

- Yan, S., Tang, C.-H., & Fu, T.-C. (2008). An airline scheduling model and solution algorithms under stochastic demands. *European Journal of Operational Research*, 190(1), 22–39. <http://doi.org/https://doi.org/10.1016/j.ejor.2007.05.053>
- Yannakakis, G. N., & Togelius, J. (2017). Artificial Intelligence and Games (First Public Draft), 359. Retrieved from <http://gameaibook.org/book.pdf>

A. Anexo: Código en GAMS del caso básico de estudio

*Definición de conjuntos

set F tipos de avion /A320,B757/ ;

set L vuelos originales /001*036/ ;

alias (l,i,j);

set LF(l) vuelos obligatorios /001, 002, 003, 006, 007, 008, 009, 011, 012, 013, 014, 016, 017, 018, 019, 020, 021, 022, 024, 025, 026, 027, 029, 030, 031, 032, 035, 036 / ;

set LO(l) vuelos opcionales /004, 005, 010, 015, 023, 028, 033, 034 / ;

set V conjunto ampliado de vuelos (originales y copias) /001_-1, 001_0, 001_1, 002_-1, 002_0, 002_1, 003_-1, 003_0, 003_1, 004_-1, 004_0, 004_1, 005_-1, 005_0, 005_1, 006_-1, 006_0, 006_1, 007_-1, 007_0, 007_1, 008_-1, 008_0, 008_1, 009_-1, 009_0, 009_1, 010_-1, 010_0, 010_1, 011_-1, 011_0, 011_1, 012_-1, 012_0, 012_1, 013_-1, 013_0, 013_1, 014_-1, 014_0, 014_1, 015_-1, 015_0, 015_1, 016_-1, 016_0, 016_1, 017_-1, 017_0, 017_1, 018_-1, 018_0, 018_1, 019_-1, 019_0, 019_1, 020_-1, 020_0, 020_1, 021_-1, 021_0, 021_1, 022_-1, 022_0, 022_1, 023_-1, 023_0, 023_1, 024_-1, 024_0, 024_1, 025_-1, 025_0, 025_1, 026_-1, 026_0, 026_1, 027_-1, 027_0, 027_1, 028_-1, 028_0, 028_1, 029_-1, 029_0, 029_1, 030_-1, 030_0, 030_1, 031_-1, 031_0, 031_1, 032_-1, 032_0, 032_1, 033_-1, 033_0, 033_1, 034_-1, 034_0, 034_1, 035_-1, 035_0, 035_1, 036_-1, 036_0, 036_1 /;

set N nodos de la red /001_-1_out, 001_0_out, 013_-1_out, 001_1_out, 004_-1_in, 013_0_out, 004_0_in, 007_-1_out, 013_1_out, 016_-1_in, 002_-1_out, 004_1_in, 007_0_out, 016_0_in, 002_0_out, 007_1_out, 016_1_in, 017_-1_in, 002_1_out, 010_-1_in, 014_-1_out, 017_0_in, 008_-1_out, 010_0_in, 014_0_out, 017_1_in, 005_-1_in, 008_0_out, 010_1_in, 014_1_out, 005_0_in, 008_1_out, 015_-1_out, 005_1_in, 009_-1_out, 011_-1_in, 015_0_out, 006_-1_in, 009_0_out, 011_0_in, 015_1_out, 006_0_in, 009_1_out, 011_1_in, 003_-1_out, 006_1_in, 018_-1_in, 003_0_out, 018_0_in, 003_1_out, 018_1_in, 012_-1_in, 012_0_in, 012_1_in, 019_-1_out, 019_0_out, 025_-1_out, 001_-1_in, 004_-1_out, 019_1_out, 025_0_out, 028_-1_in, 001_0_in, 004_0_out, 025_1_out, 028_0_in, 001_1_in, 004_1_out, 028_1_in, 026_-1_out, 002_-1_in, 020_-1_out, 022_-1_in, 026_0_out, 029_-1_in, 002_0_in, 020_0_out, 022_0_in, 026_1_out, 029_0_in, 002_1_in, 005_-1_out, 020_1_out, 022_1_in, 029_1_in, 005_0_out, 023_-1_in, 005_1_out, 023_0_in, 006_-1_out, 023_1_in, 006_0_out, 027_-1_out, 006_1_out, 021_-1_out, 027_0_out, 021_0_out, 027_1_out, 003_-1_in, 021_1_out, 024_-1_in, 030_-1_in, 003_0_in, 024_0_in, 030_0_in, 003_1_in, 024_1_in, 030_1_in, 028_-1_out, 025_-1_in, 028_0_out, 013_-1_in, 016_-1_out, 025_0_in, 028_1_out, 031_-1_out, 013_0_in, 016_0_out, 025_1_in, 031_0_out, 013_1_in, 016_1_out, 017_-1_out, 029_-1_out, 031_1_out, 017_0_out, 026_-1_in, 029_0_out, 034_-1_in, 017_1_out, 026_0_in, 029_1_out, 032_-1_out, 034_0_in, 014_-1_in, 026_1_in, 032_0_out, 034_1_in, 014_0_in, 032_1_out, 014_1_in, 033_-1_out, 035_-1_in, 015_-1_in, 033_0_out, 035_0_in, 036_-1_in, 015_0_in, 033_1_out, 035_1_in, 036_0_in, 015_1_in, 018_-1_out, 027_-1_in, 036_1_in, 018_0_out, 027_0_in, 030_-1_out, 018_1_out, 027_1_in, 030_0_out, 030_1_out, 019_-1_in, 019_0_in, 019_1_in, 031_-1_in, 007_-1_in, 022_-1_out, 031_0_in, 034_-1_out, 007_0_in, 010_-1_out, 022_0_out, 031_1_in, 034_0_out, 007_1_in, 010_0_out, 020_-1_in, 022_1_out, 034_1_out, 010_1_out, 020_0_in, 023_-1_out, 032_-1_in, 008_-1_in, 020_1_in, 023_0_out, 032_0_in, 035_-1_out, 008_0_in, 011_-1_out, 023_1_out, 032_1_in, 035_0_out, 036_-1_out, 008_1_in, 011_0_out, 033_-1_in, 035_1_out, 036_0_out, 009_-1_in, 011_1_out, 033_0_in, 036_1_out, 009_0_in, 033_1_in, 009_1_in, 021_-1_in, 024_-1_out, 021_0_in, 024_0_out, 012_-1_out, 021_1_in, 024_1_out, 012_0_out, 012_1_out / ;

set U(N) conjunto de últimos nodos del día en cada estacion /012_1_in, 030_1_in, 030_1_out, 012_1_out / ;

*Definición de parámetros

Parameter C(f,v) costo de asignar el tipo de avión f al vuelo v /

A320.001_-1	134	A320.010_-1	295	A320.019_-1	406	A320.028_-1	175
B757.001_-1	161	B757.010_-1	354	B757.019_-1	487	B757.028_-1	210
A320.001_0	134	A320.010_0	295	A320.019_0	406	A320.028_0	175
B757.001_0	161	B757.010_0	354	B757.019_0	487	B757.028_0	210
A320.001_1	134	A320.010_1	295	A320.019_1	406	A320.028_1	175
B757.001_1	161	B757.010_1	354	B757.019_1	487	B757.028_1	210
A320.002_-1	134	A320.011_-1	295	A320.020_-1	406	A320.029_-1	175
B757.002_-1	161	B757.011_-1	354	B757.020_-1	487	B757.029_-1	210
A320.002_0	134	A320.011_0	295	A320.020_0	406	A320.029_0	175
B757.002_0	161	B757.011_0	354	B757.020_0	487	B757.029_0	210
A320.002_1	134	A320.011_1	295	A320.020_1	406	A320.029_1	175
B757.002_1	161	B757.011_1	354	B757.020_1	487	B757.029_1	210
A320.003_-1	134	A320.012_-1	295	A320.021_-1	406	A320.030_-1	175
B757.003_-1	161	B757.012_-1	354	B757.021_-1	487	B757.030_-1	210
A320.003_0	134	A320.012_0	295	A320.021_0	406	A320.030_0	175
B757.003_0	161	B757.012_0	354	B757.021_0	487	B757.030_0	210
A320.003_1	134	A320.012_1	295	A320.021_1	406	A320.030_1	175
B757.003_1	161	B757.012_1	354	B757.021_1	487	B757.030_1	210
A320.004_-1	134	A320.013_-1	192	A320.022_-1	406	A320.031_-1	478
B757.004_-1	161	B757.013_-1	230	B757.022_-1	487	B757.031_-1	574
A320.004_0	134	A320.013_0	192	A320.022_0	406	A320.031_0	478
B757.004_0	161	B757.013_0	230	B757.022_0	487	B757.031_0	574
A320.004_1	134	A320.013_1	192	A320.022_1	406	A320.031_1	478
B757.004_1	161	B757.013_1	230	B757.022_1	487	B757.031_1	574
A320.005_-1	134	A320.014_-1	192	A320.023_-1	406	A320.032_-1	478
B757.005_-1	161	B757.014_-1	230	B757.023_-1	487	B757.032_-1	574
A320.005_0	134	A320.014_0	192	A320.023_0	406	A320.032_0	478
B757.005_0	161	B757.014_0	230	B757.023_0	487	B757.032_0	574
A320.005_1	134	A320.014_1	192	A320.023_1	406	A320.032_1	478
B757.005_1	161	B757.014_1	230	B757.023_1	487	B757.032_1	574
A320.006_-1	134	A320.015_-1	192	A320.024_-1	406	A320.033_-1	478
B757.006_-1	161	B757.015_-1	230	B757.024_-1	487	B757.033_-1	574
A320.006_0	134	A320.015_0	192	A320.024_0	406	A320.033_0	478
B757.006_0	161	B757.015_0	230	B757.024_0	487	B757.033_0	574
A320.006_1	134	A320.015_1	192	A320.024_1	406	A320.033_1	478
B757.006_1	161	B757.015_1	230	B757.024_1	487	B757.033_1	574
A320.007_-1	295	A320.016_-1	192	A320.025_-1	175	A320.034_-1	478
B757.007_-1	354	B757.016_-1	230	B757.025_-1	210	B757.034_-1	574
A320.007_0	295	A320.016_0	192	A320.025_0	175	A320.034_0	478
B757.007_0	354	B757.016_0	230	B757.025_0	210	B757.034_0	574
A320.007_1	295	A320.016_1	192	A320.025_1	175	A320.034_1	478
B757.007_1	354	B757.016_1	230	B757.025_1	210	B757.034_1	574
A320.008_-1	295	A320.017_-1	192	A320.026_-1	175	A320.035_-1	478
B757.008_-1	354	B757.017_-1	230	B757.026_-1	210	B757.035_-1	574
A320.008_0	295	A320.017_0	192	A320.026_0	175	A320.035_0	478
B757.008_0	354	B757.017_0	230	B757.026_0	210	B757.035_0	574
A320.008_1	295	A320.017_1	192	A320.026_1	175	A320.035_1	478
B757.008_1	354	B757.017_1	230	B757.026_1	210	B757.035_1	574
A320.009_-1	295	A320.018_-1	192	A320.027_-1	175	A320.036_-1	478
B757.009_-1	354	B757.018_-1	230	B757.027_-1	210	B757.036_-1	574
A320.009_0	295	A320.018_0	192	A320.027_0	175	A320.036_0	478
B757.009_0	354	B757.018_0	230	B757.027_0	210	B757.036_0	574
A320.009_1	295	A320.018_1	192	A320.027_1	175	A320.036_1	478
B757.009_1	354	B757.018_1	230	B757.027_1	210	B757.036_1	574

/;

parameter A(f) número de aviones disponibles del tipo f /

A320 6
B757 4 /;

parameter r(l) pérdida de rentabilidad si el vuelo l se excluye de la red /

001	168	013	240	025	219
002	168	014	240	026	219
003	168	015	240	027	219
004	168	016	240	028	219
005	168	017	240	029	219
006	168	018	240	030	219
007	369	019	508	031	598
008	369	020	508	032	598
009	369	021	508	033	598
010	369	022	508	034	598
011	369	023	508	035	598
012	369	024	508	036	598

/;

parameter Conect(v,n) 1 si el vuelo l es una llegada al nodo n -1 si el vuelo l es una salida del nodo n 0 en caso contrario /

001_-1.001_-1_out	-1	007_1.007_1_out	-1	014_0.014_0_out	-1
001_-1.001_-1_in	1	007_1.007_1_in	1	014_0.014_0_in	1
001_0.001_0_out	-1	008_-1.008_-1_out	-1	014_1.014_1_out	-1
001_0.001_0_in	1	008_-1.008_-1_in	1	014_1.014_1_in	1
001_1.001_1_out	-1	008_0.008_0_out	-1	015_-1.015_-1_out	-1
001_1.001_1_in	1	008_0.008_0_in	1	015_-1.015_-1_in	1
002_-1.002_-1_out	-1	008_1.008_1_out	-1	015_0.015_0_out	-1
002_-1.002_-1_in	1	008_1.008_1_in	1	015_0.015_0_in	1
002_0.002_0_out	-1	009_-1.009_-1_out	-1	015_1.015_1_out	-1
002_0.002_0_in	1	009_-1.009_-1_in	1	015_1.015_1_in	1
002_1.002_1_out	-1	009_0.009_0_out	-1	016_-1.016_-1_in	1
002_1.002_1_in	1	009_0.009_0_in	1	016_-1.016_-1_out	-1
003_-1.003_-1_out	-1	009_1.009_1_out	-1	016_0.016_0_in	1
003_-1.003_-1_in	1	009_1.009_1_in	1	016_0.016_0_out	-1
003_0.003_0_out	-1	010_-1.010_-1_in	1	016_1.016_1_in	1
003_0.003_0_in	1	010_-1.010_-1_out	-1	016_1.016_1_out	-1
003_1.003_1_out	-1	010_0.010_0_in	1	017_-1.017_-1_in	1
003_1.003_1_in	1	010_0.010_0_out	-1	017_-1.017_-1_out	-1
004_-1.004_-1_in	1	010_1.010_1_in	1	017_0.017_0_in	1
004_-1.004_-1_out	-1	010_1.010_1_out	-1	017_0.017_0_out	-1
004_0.004_0_in	1	011_-1.011_-1_in	1	017_1.017_1_in	1
004_0.004_0_out	-1	011_-1.011_-1_out	-1	017_1.017_1_out	-1
004_1.004_1_in	1	011_0.011_0_in	1	018_-1.018_-1_in	1
004_1.004_1_out	-1	011_0.011_0_out	-1	018_-1.018_-1_out	-1
005_-1.005_-1_in	1	011_1.011_1_in	1	018_0.018_0_in	1
005_-1.005_-1_out	-1	011_1.011_1_out	-1	018_0.018_0_out	-1
005_0.005_0_in	1	012_-1.012_-1_in	1	018_1.018_1_in	1
005_0.005_0_out	-1	012_-1.012_-1_out	-1	018_1.018_1_out	-1
005_1.005_1_in	1	012_0.012_0_in	1	019_-1.019_-1_out	-1
005_1.005_1_out	-1	012_0.012_0_out	-1	019_-1.019_-1_in	1
006_-1.006_-1_in	1	012_1.012_1_in	1	019_0.019_0_out	-1
006_-1.006_-1_out	-1	012_1.012_1_out	-1	019_0.019_0_in	1
006_0.006_0_in	1	013_-1.013_-1_out	-1	019_1.019_1_out	-1
006_0.006_0_out	-1	013_-1.013_-1_in	1	019_1.019_1_in	1
006_1.006_1_in	1	013_0.013_0_out	-1	020_-1.020_-1_out	-1
006_1.006_1_out	-1	013_0.013_0_in	1	020_-1.020_-1_in	1
007_-1.007_-1_out	-1	013_1.013_1_out	-1	020_0.020_0_out	-1
007_-1.007_-1_in	1	013_1.013_1_in	1	020_0.020_0_in	1
007_0.007_0_out	-1	014_-1.014_-1_out	-1	020_1.020_1_out	-1
007_0.007_0_in	1	014_-1.014_-1_in	1	020_1.020_1_in	1

```

021_-1.021_-1_out -1      026_0.026_0_out -1      031_1.031_1_out -1
021_-1.021_-1_in  1       026_0.026_0_in  1       031_1.031_1_in  1
021_0.021_0_out -1       026_1.026_1_out -1      032_-1.032_-1_out -1
021_0.021_0_in  1       026_1.026_1_in  1       032_-1.032_-1_in  1
021_1.021_1_out -1       027_-1.027_-1_out -1      032_0.032_0_out -1
021_1.021_1_in  1       027_-1.027_-1_in  1       032_0.032_0_in  1
022_-1.022_-1_in  1       027_0.027_0_out -1      032_1.032_1_out -1
022_-1.022_-1_out -1     027_0.027_0_in  1       032_1.032_1_in  1
022_0.022_0_in  1       027_1.027_1_out -1      033_-1.033_-1_out -1
022_0.022_0_out -1      027_1.027_1_in  1       033_-1.033_-1_in  1
022_1.022_1_in  1       028_-1.028_-1_in  1       033_0.033_0_out -1
022_1.022_1_out -1      028_-1.028_-1_out -1      033_0.033_0_in  1
023_-1.023_-1_in  1       028_0.028_0_in  1       033_1.033_1_out -1
023_-1.023_-1_out -1     028_0.028_0_out -1      033_1.033_1_in  1
023_0.023_0_in  1       028_1.028_1_in  1       034_-1.034_-1_in  1
023_0.023_0_out -1      028_1.028_1_out -1      034_-1.034_-1_out -1
023_1.023_1_in  1       029_-1.029_-1_in  1       034_0.034_0_in  1
023_1.023_1_out -1      029_-1.029_-1_out -1      034_0.034_0_out -1
024_-1.024_-1_in  1       029_0.029_0_in  1       034_1.034_1_in  1
024_-1.024_-1_out -1     029_0.029_0_out -1      034_1.034_1_out -1
024_0.024_0_in  1       029_1.029_1_in  1       035_-1.035_-1_in  1
024_0.024_0_out -1      029_1.029_1_out -1      035_-1.035_-1_out -1
024_1.024_1_in  1       030_-1.030_-1_in  1       035_0.035_0_in  1
024_1.024_1_out -1      030_-1.030_-1_out -1      035_0.035_0_out -1
025_-1.025_-1_out -1     030_0.030_0_in  1       035_1.035_1_in  1
025_-1.025_-1_in  1       030_0.030_0_out -1      035_1.035_1_out -1
025_0.025_0_out -1      030_1.030_1_in  1       036_-1.036_-1_in  1
025_0.025_0_in  1       030_1.030_1_out -1      036_-1.036_-1_out -1
025_1.025_1_out -1      031_-1.031_-1_out -1      036_0.036_0_in  1
025_1.025_1_in  1       031_-1.031_-1_in  1       036_0.036_0_out -1
026_-1.026_-1_out -1     031_0.031_0_out -1      036_1.036_1_in  1
026_-1.026_-1_in  1       031_0.031_0_in  1       036_1.036_1_out -1
/;

```

parameter K(l,v) 1 si el vuelo v es una copia del vuelo original l 0 en caso contrario /

```

001.001_-1  1      009.009_1  1      018.018_0  1
001.001_0   1      010.010_-1  1      018.018_1  1
001.001_1   1      010.010_0  1      019.019_-1  1
002.002_-1  1      010.010_1  1      019.019_0  1
002.002_0   1      011.011_-1  1      019.019_1  1
002.002_1   1      011.011_0  1      020.020_-1  1
003.003_-1  1      011.011_1  1      020.020_0  1
003.003_0   1      012.012_-1  1      020.020_1  1
003.003_1   1      012.012_0  1      021.021_-1  1
004.004_-1  1      012.012_1  1      021.021_0  1
004.004_0   1      013.013_-1  1      021.021_1  1
004.004_1   1      013.013_0  1      022.022_-1  1
005.005_-1  1      013.013_1  1      022.022_0  1
005.005_0   1      014.014_-1  1      022.022_1  1
005.005_1   1      014.014_0  1      023.023_-1  1
006.006_-1  1      014.014_1  1      023.023_0  1
006.006_0   1      015.015_-1  1      023.023_1  1
006.006_1   1      015.015_0  1      024.024_-1  1
007.007_-1  1      015.015_1  1      024.024_0  1
007.007_0   1      016.016_-1  1      024.024_1  1
007.007_1   1      016.016_0  1      025.025_-1  1
008.008_-1  1      016.016_1  1      025.025_0  1
008.008_0   1      017.017_-1  1      025.025_1  1
008.008_1   1      017.017_0  1      026.026_-1  1
009.009_-1  1      017.017_1  1      026.026_0  1
009.009_0   1      018.018_-1  1      026.026_1  1

```

```

027.027_-1 1          030.030_0 1          033.033_1 1
027.027_0 1          030.030_1 1          034.034_-1 1
027.027_1 1          031.031_-1 1         034.034_0 1
028.028_-1 1          031.031_0 1         034.034_1 1
028.028_0 1          031.031_1 1         035.035_-1 1
028.028_1 1          032.032_-1 1         035.035_0 1
029.029_-1 1          032.032_0 1         035.035_1 1
029.029_0 1          032.032_1 1         036.036_-1 1
029.029_1 1          033.033_-1 1         036.036_0 1
030.030_-1 1          033.033_0 1         036.036_1 1
/;

```

parameter HS(v) hora de salida del vuelo v /

```

001_-1 5          010_-1 10          019_-1 4          028_-1 5
001_0 6           010_0 11           019_0 5          028_0 6
001_1 7           010_1 12           019_1 6          028_1 7
002_-1 9          011_-1 14          020_-1 10         029_-1 9
002_0 10          011_0 15          020_0 11         029_0 10
002_1 11          011_1 16          020_1 12         029_1 11
003_-1 18         012_-1 20          021_-1 17         030_-1 18
003_0 19          012_0 21          021_0 18         030_0 19
003_1 20          012_1 22          021_1 19         030_1 20
004_-1 6          013_-1 6           022_-1 9          031_-1 7
004_0 7           013_0 7           022_0 10         031_0 8
004_1 8           013_1 8           022_1 11         031_1 9
005_-1 12         014_-1 11          023_-1 12         032_-1 11
005_0 13          014_0 12          023_0 13         032_0 12
005_1 14          014_1 13          023_1 14         032_1 13
006_-1 15         015_-1 14          024_-1 18         033_-1 14
006_0 16          015_0 15          024_0 19         033_0 15
006_1 17          015_1 16          024_1 20         033_1 16
007_-1 8          016_-1 7           025_-1 5          034_-1 9
007_0 9           016_0 8           025_0 6          034_0 10
007_1 10          016_1 9           025_1 7          034_1 11
008_-1 12         017_-1 9           026_-1 9          035_-1 13
008_0 13          017_0 10          026_0 10         035_0 14
008_1 14          017_1 11          026_1 11         035_1 15
009_-1 15         018_-1 17          027_-1 16         036_-1 14
009_0 16          018_0 18          027_0 17         036_0 15
009_1 17          018_1 19          027_1 18         036_1 16
/;

```

parameter HL(v) hora de llegada del vuelo v /

```

001_-1 6          007_0 10          013_1 9          020_-1 11
001_0 7           007_1 11          014_-1 12         020_0 12
001_1 8           008_-1 13         014_0 13         020_1 13
002_-1 10         008_0 14          014_1 14         021_-1 18
002_0 11          008_1 15          015_-1 15         021_0 19
002_1 12          009_-1 16         015_0 16         021_1 20
003_-1 19         009_0 17          015_1 17         022_-1 10
003_0 20          009_1 18          016_-1 8         022_0 11
003_1 21          010_-1 11         016_0 9          022_1 12
004_-1 7          010_0 12          016_1 10         023_-1 13
004_0 8           010_1 13          017_-1 10         023_0 14
004_1 9           011_-1 15         017_0 11         023_1 15
005_-1 13         011_0 16          017_1 12         024_-1 19
005_0 14          011_1 17          018_-1 18         024_0 20
005_1 15          012_-1 21         018_0 19         024_1 21
006_-1 16         012_0 22          018_1 20         025_-1 6
006_0 17          012_1 23          019_-1 5         025_0 7
006_1 18          013_-1 7           019_0 6         025_1 8
007_-1 9          013_0 8           019_1 7         026_-1 10

```

026_0	11	029_-1	10	031_1	10	034_0	11
026_1	12	029_0	11	032_-1	12	034_1	12
027_-1	17	029_1	12	032_0	13	035_-1	14
027_0	18	030_-1	19	032_1	14	035_0	15
027_1	19	030_0	20	033_-1	15	035_1	16
028_-1	6	030_1	21	033_0	16	036_-1	15
028_0	7	031_-1	8	033_1	17	036_0	16
028_1	8	031_0	9	034_-1	10	036_1	17

/;

*Para las salidas y llegadas se usará la siguiente notación

*MDE 1

*BOG 2

*CLO 3

*CTG 4

parameter salida(l) aeropuerto de salida del vuelo l /

001	1	010	4	019	2	028	3
002	1	011	4	020	2	029	3
003	1	012	4	021	2	030	3
004	2	013	1	022	4	031	3
005	2	014	1	023	4	032	3
006	2	015	1	024	4	033	3
007	1	016	3	025	2	034	4
008	1	017	3	026	2	035	4
009	1	018	3	027	2	036	4

/;

parameter llegada(l) aeropuerto de llegada del vuelo l /

001	2	010	1	019	4	028	2
002	2	011	1	020	4	029	2
003	2	012	1	021	4	030	2
004	1	013	3	022	2	031	4
005	1	014	3	023	2	032	4
006	1	015	3	024	2	033	4
007	4	016	1	025	3	034	3
008	4	017	1	026	3	035	3
009	4	018	1	027	3	036	3

/;

scalar rho tiempo minimo para que una conexión entre dos vuelos sea factible /0.5/ ;

variables

X(f,v) 1 si el avion de tipo f se asigna al vuelo v 0 en caso contrario

Y(f,n) aviones de tipo f en tierra en el nodo n

Z(i,j) 1 si los vuelos i-j originales tienen alguna conexión factible en alguna de sus copias 0 en caso contrario
valfo valor de la función objetivo

valcostoreal valor del costo sin el componente de penalización

numce numero total de conexiones exitosas;

binary variables X,Z ;

integer variables Y ;

*Se fija como 0 la variable z para vuelos que no llegan y salen del mismo aeropuerto

z.fx(i,j)\$(salida(j)<>llegada(i))=0;

*El límite superior de Y se fija muy alto (en comparación con el tamaño del problema) para no tener problemas con la licencia

*Las licencias demo de GAMS normalmente fijan los valores de las variables enteras.


```
y.up(f,n)=80 ;
```

```
scalar M1 penalizacion unitaria por conexion infactible /125/ ;
scalar M2 número muy grande para las restricciones que identifican si una conexion es factible o no /100000/ ;
```

```
equations
```

```
fo valor de la funcion objetivo
```

```
costoreal valor de la funcion objetivo sin el componente de penalizacion
```

```
totalconex numero total de conexiones exitosas asignadas
```

```
balance ecuaciones de balance siempre que no sea primer y ultimo nodo de cada aeropuerto
```

```
balanceMDE balance en los primeros y últimos nodos de MDE
```

```
balanceBOG balance en los primeros y últimos nodos de BOG
```

```
balanceCLO balance en los primeros y últimos nodos de CLO
```

```
balanceCTG balance en los primeros y últimos nodos de CTG
```

```
disponibilidad disponibilidad de cada tipo de avion
```

```
cobobli cobertura de vuelos obligatorios
```

```
cobopt cobertura de los vuelos opcionales
```

```
relz relacion 1 entre la variable z y las variables x
```

```
defz z puede ser 1 si los vuelos son 1;
```

```
fo .. valfo =e= sum((f,v),C(f,v)*X(f,v)) + sum(l,r(l)*(1-sum((f,v),K(l,v)*X(f,v)))) +
M1*sum((i,j)$(salida(j)=llegada(i)),(1-z(i,j)));
costoreal .. valcostoreal =e= sum((f,v),C(f,v)*X(f,v)) + sum(l,r(l)*(1-sum((f,v),K(l,v)*X(f,v)))) ;
totalconex .. numce =e= sum((i,j)$(salida(j)=llegada(i)),z(i,j));
balance(f,n)$(ord(n)<>1 and ord(n)<>55 and ord(n)<>109 and ord(n)<>163) .. y(f,n) =e= y(f,n-1)
+sum(v,Conect(v,n)*X(f,v));
balanceMDE(f,n)$(ord(n)=1) .. y(f,n) =e= y(f,n+53) + sum(v,Conect(v,n)*X(f,v));
balanceBOG(f,n)$(ord(n)=55) .. y(f,n) =e= y(f,n+53) + sum(v,Conect(v,n)*X(f,v));
balanceCLO(f,n)$(ord(n)=109) .. y(f,n) =e= y(f,n+53) + sum(v,Conect(v,n)*X(f,v));
balanceCTG(f,n)$(ord(n)=163) .. y(f,n) =e= y(f,n+53) + sum(v,Conect(v,n)*X(f,v));
disponibilidad(f) .. sum(u,y(f,u)) =l= A(f) ;
cobobli(lf) .. sum((f,v),K(lf,v)*X(f,v)) =e= 1 ;
cobopt(lo) .. sum((f,v),K(lo,v)*X(f,v)) =l= 1 ;
relz(i,j)$(salida(j)=llegada(i)) .. z(i,j) =l= (sum((f,v),K(j,v)*HS(v)*X(f,v))-sum((f,v),K(i,v)*HL(v)*X(f,v)))/M2 + 1 ;
defz(i,j)$(salida(j)=llegada(i)) .. 2*z(i,j) =l= sum((f,v),K(i,v)*X(f,v)) + sum((f,v),K(j,v)*X(f,v)) ;
```

```
*Opción para que muestre todas las ecuaciones del modelo
```

```
option limrow=30000;
```

```
*Límite de iteraciones máximo
```

```
*option iterlim=1000;
```

```
model integratedassignment /all/ ;
```

```
solve integratedassignment using mip minimizing valfo ;
```

```
*Opción para mostrar los valores de la variable X en forma de lista
```

```
option x:2:0:1; display x.l;
```

```
display valfo.l, costoreal.l, numce.l ;
```


B. Anexo: Itinerario base y parámetros del caso de aplicación

Itinerario base

FN	Origen	Destino	Salida	Llegada	Obligatorio
AV9779	ADZ	BOG	14,08	16,08	No
AV9778	BOG	ADZ	11,50	13,58	No
AV9777	ADZ	BOG	16,83	18,83	No
AV9776	BOG	ADZ	14,17	16,25	No
AV9781	ADZ	BOG	22,00	0,17	No
AV9780	BOG	ADZ	19,25	21,55	No
AV9001	ADZ	CTG	12,08	13,58	No
AV9002	CTG	ADZ	10,17	11,67	No
AV9004	ADZ	CTG	13,33	14,83	No
AV9003	CTG	ADZ	11,42	12,92	No
AV91	AUA	BOG	18,67	19,58	Sí
AV90	BOG	AUA	14,33	17,17	Sí
AV9940	AXM	BOG	7,35	8,25	No
AV9939	BOG	AXM	6,00	7,10	No
AV9938	AXM	BOG	9,83	10,73	No
AV9937	BOG	AXM	8,58	9,58	No
AV9942	AXM	BOG	12,33	13,23	No
AV9941	BOG	AXM	11,08	12,03	No
AV9944	AXM	BOG	16,80	17,72	No
AV9943	BOG	AXM	15,42	16,47	No
AV9936	AXM	BOG	19,50	20,47	No
AV9935	BOG	AXM	18,25	19,25	No
AV9521	BAQ	BOG	6,08	7,58	Sí
AV8522	BOG	BAQ	5,78	7,42	Sí
AV8533	BAQ	BOG	7,00	8,50	Sí
AV9536	BOG	BAQ	6,25	7,83	Sí
AV9005	BAQ	BOG	7,50	9,00	No
AV9522	BOG	BAQ	7,50	9,08	No
AV8523	BAQ	BOG	7,83	9,33	Sí

AV9534	BOG	BAQ	8,50	10,08	Sí
AV9537	BAQ	BOG	8,33	9,75	Sí
AV8520	BOG	BAQ	9,75	11,25	No
AV9006	BAQ	BOG	8,33	9,83	Sí
AV9007	BOG	BAQ	10,42	11,92	Sí
AV9533	BAQ	BOG	10,67	12,08	No
AV9538	BOG	BAQ	11,92	13,33	Sí
AV8521	BAQ	BOG	11,75	13,17	Sí
AV8528	BOG	BAQ	15,25	16,83	Sí
AV9008	BAQ	BOG	12,42	14,00	Sí
AV8532	BOG	BAQ	17,42	19,00	No
AV9539	BAQ	BOG	13,83	15,25	Sí
AV9530	BOG	BAQ	19,67	21,08	Sí
AV8529	BAQ	BOG	17,33	18,80	No
AV9532	BOG	BAQ	20,83	22,48	Sí
AV9535	BAQ	BOG	17,95	19,37	No
AV8530	BOG	BAQ	21,92	23,57	No
AV8531	BAQ	BOG	19,42	20,88	Sí
AV9531	BAQ	BOG	21,58	23,00	Sí
AV9483	BAQ	MDE	7,83	9,45	No
AV9482	MDE	BAQ	6,00	7,52	No
AV9369	BAQ	MDE	18,17	19,87	No
AV9368	MDE	BAQ	16,42	17,92	No
AV2	BAQ	MIA	10,17	13,92	Sí
AV3	MIA	BAQ	15,17	16,78	Sí
AV19	BCN	BOG	16,67	21,08	Sí
AV18	BOG	BCN	21,42	14,83	Sí
AV9471	BGA	BOG	6,00	7,00	No
AV8572	BOG	BGA	6,17	7,32	No
AV8447	BOG	BGA	7,75	8,72	No

AV9009	BOG	BGA	7,42	8,42	No
AV9010	BOG	BGA	9,00	9,92	No
AV9460	BOG	BGA	7,42	8,57	No
AV9461	BOG	BGA	9,00	9,97	No
AV9472	BOG	BGA	9,17	10,23	No
AV9473	BOG	BGA	10,73	11,73	No
AV9468	BOG	BGA	10,50	11,50	No
AV9469	BOG	BGA	11,92	12,83	No
AV8578	BOG	BGA	12,50	13,50	No
AV8579	BOG	BGA	13,92	14,92	No
AV9464	BOG	BGA	14,92	15,92	No
AV9465	BOG	BGA	16,50	17,42	No
AV8564	BOG	BGA	15,08	16,08	No
AV8565	BOG	BGA	16,58	17,58	No
AV9462	BOG	BGA	17,17	18,17	No
AV9463	BOG	BGA	18,58	19,58	No
AV8566	BOG	BGA	18,42	19,42	No
AV8567	BOG	BGA	19,83	20,83	No
AV9466	BOG	BGA	19,50	20,50	No
AV9467	BOG	BGA	21,00	21,92	No
AV9470	BOG	BGA	21,92	23,00	No
AV66	BOG	CCS	8,50	10,83	Sí
AV79	CCS	BOG	6,25	7,58	Sí
AV9011	BOG	CCS	8,50	11,00	Sí
AV69	CCS	BOG	11,75	13,25	Sí
AV80	BOG	CCS	14,42	16,75	Sí
AV9012	CCS	BOG	11,92	13,25	Sí
AV78	BOG	CCS	21,67	0,08	Sí
AV81	CCS	BOG	18,67	20,17	Sí
AV9201	BOG	CLO	6,00	7,08	Sí
AV9200	CLO	BOG	6,00	6,87	Sí
AV9217	BOG	CLO	6,25	7,33	No
AV9236	CLO	BOG	6,50	7,37	Sí
AV8401	BOG	CLO	6,50	7,67	Sí
AV9220	CLO	BOG	7,00	7,87	Sí
AV9203	BOG	CLO	6,92	8,00	Sí
AV9202	CLO	BOG	7,58	8,45	Sí
AV9231	BOG	CLO	8,00	9,08	Sí
AV8412	CLO	BOG	7,58	8,50	Sí
AV9207	BOG	CLO	8,92	10,00	No

AV8400	CLO	BOG	8,08	9,00	Sí
AV9209	BOG	CLO	10,00	11,08	Sí
AV9218	CLO	BOG	8,33	9,20	No
AV9204	CLO	BOG	8,50	9,37	Sí
AV9211	BOG	CLO	11,00	12,08	Sí
AV8418	CLO	BOG	9,58	10,50	No
AV8409	BOG	CLO	12,00	13,17	No
AV9208	CLO	BOG	10,58	11,45	Sí
AV9221	BOG	CLO	13,00	14,08	Sí
AV9205	BOG	CLO	14,00	15,08	Sí
AV9210	CLO	BOG	11,67	12,53	No
AV9212	CLO	BOG	12,58	13,45	Sí
AV8443	BOG	CLO	14,00	15,17	Sí
AV8420	CLO	BOG	13,58	14,58	Sí
AV8421	BOG	CLO	15,00	16,17	No
AV9222	CLO	BOG	14,58	15,45	Sí
AV9013	BOG	CLO	15,58	16,67	Sí
AV9206	CLO	BOG	15,58	16,45	No
AV9213	BOG	CLO	15,92	17,00	No
AV9215	BOG	CLO	17,00	18,08	Sí
AV8442	CLO	BOG	15,58	16,58	Sí
AV9227	BOG	CLO	18,00	19,08	No
AV9226	CLO	BOG	16,12	16,98	Sí
AV8405	BOG	CLO	18,50	19,67	Sí
AV8422	CLO	BOG	16,67	17,67	Sí
AV9233	BOG	CLO	19,08	20,25	Sí
AV9214	CLO	BOG	17,58	18,45	No
AV9223	BOG	CLO	20,00	21,08	Sí
AV9014	CLO	BOG	18,58	19,50	Sí
AV9219	BOG	CLO	20,50	21,58	Sí
AV9216	CLO	BOG	19,00	19,97	Sí
AV9015	BOG	CLO	21,17	22,25	Sí
AV8406	CLO	BOG	20,08	21,08	Sí
AV8417	BOG	CLO	22,00	23,17	Sí
AV9234	CLO	BOG	20,67	21,67	No
AV9224	CLO	BOG	21,58	22,55	Sí
AV9540	BOG	CTG	7,42	8,83	Sí
AV9755	CTG	BOG	6,00	7,33	Sí
AV9786	BOG	CTG	8,17	9,67	Sí
AV9541	CTG	BOG	9,33	10,67	Sí

AV9542	BOG	CTG	8,75	10,17	Sí
AV9543	CTG	BOG	10,67	12,00	No
AV8490	BOG	CTG	9,42	10,92	Sí
AV9549	CTG	BOG	11,42	12,75	Sí
AV9548	BOG	CTG	9,50	10,92	Sí
AV9018	CTG	BOG	12,17	13,50	No
AV9019	BOG	CTG	10,25	11,67	Sí
AV9759	CTG	BOG	12,25	13,58	No
AV9758	BOG	CTG	10,33	11,75	No
AV9753	CTG	BOG	14,00	15,33	No
AV9752	BOG	CTG	12,08	13,50	Sí
AV9787	CTG	BOG	14,00	15,50	Sí
AV9788	BOG	CTG	13,00	14,50	Sí
AV9789	CTG	BOG	14,92	16,42	Sí
AV9546	BOG	CTG	13,67	15,08	Sí
AV8491	CTG	BOG	15,25	16,75	Sí
AV9544	BOG	CTG	15,83	17,25	Sí
AV9547	CTG	BOG	15,58	16,92	No
AV8548	BOG	CTG	18,08	19,58	Sí
AV9545	CTG	BOG	17,75	19,08	No
AV8550	BOG	CTG	18,92	20,42	Sí
AV8549	CTG	BOG	20,00	21,38	No
AV8552	BOG	CTG	20,42	21,92	Sí
AV8551	CTG	BOG	20,83	22,25	Sí
AV9754	BOG	CTG	21,92	23,42	No
AV8553	CTG	BOG	22,33	23,83	Sí
AV9450	BOG	CUC	6,08	7,25	No
AV9451	CUC	BOG	7,75	8,83	No
AV9448	BOG	CUC	9,17	10,33	No
AV9449	CUC	BOG	10,92	12,00	No
AV9020	BOG	CUC	12,25	13,83	No
AV9021	CUC	BOG	14,08	15,58	No
AV9456	BOG	CUC	13,92	15,50	No
AV9457	CUC	BOG	15,75	17,08	No
AV9452	BOG	CUC	17,50	18,67	No
AV9453	CUC	BOG	19,17	20,25	No
AV9458	BOG	CUC	19,00	20,17	No
AV9459	CUC	BOG	20,67	21,75	No
AV92	BOG	CUR	14,67	17,67	Sí
AV93	CUR	BOG	19,00	20,00	Sí
AV8586	BOG	EJA	6,08	7,27	No

AV8587	EJA	BOG	7,68	8,60	No
AV9494	BOG	EJA	11,58	12,63	No
AV9495	EJA	BOG	13,08	14,03	No
AV9492	BOG	EJA	15,92	16,92	No
AV9493	EJA	BOG	17,42	18,45	No
AV8580	BOG	EJA	16,17	17,17	No
AV8581	EJA	BOG	17,67	18,62	No
AV87	BOG	EZE	22,17	6,42	Sí
AV88	EZE	BOG	7,75	12,25	Sí
AV36	BOG	FLL	9,73	14,50	Sí
AV37	FLL	BOG	17,33	20,00	Sí
AV85	BOG	GRU	21,60	5,52	Sí
AV86	GRU	BOG	8,58	12,58	Sí
AV113	BOG	GYE	14,67	16,50	Sí
AV280	GYE	BOG	6,17	7,92	Sí
AV281	BOG	GYE	21,75	23,50	Sí
AV112	GYE	BOG	18,00	20,00	Sí
AV286	BOG	IAD	9,35	15,60	Sí
AV287	IAD	BOG	16,60	21,02	Sí
AV9281	BOG	IBE	5,83	6,60	No
AV9282	IBE	BOG	6,92	7,58	No
AV9279	BOG	IBE	13,32	14,08	No
AV9280	IBE	BOG	14,42	15,08	No
AV9022	BOG	IBE	15,33	16,10	No
AV9023	IBE	BOG	16,42	17,08	No
AV9283	BOG	IBE	17,42	18,18	No
AV9284	IBE	BOG	18,50	19,17	No
AV284	BOG	JFK	14,72	21,50	Sí
AV285	JFK	BOG	0,92	5,50	Sí
AV20	BOG	JFK	22,92	5,50	Sí
AV21	JFK	BOG	8,00	13,00	Sí
AV48	BOG	LAX	1,58	7,27	Sí
AV49	LAX	BOG	11,50	20,75	Sí
AV25	BOG	LIM	13,67	16,75	Sí
AV74	LIM	BOG	5,83	8,77	Sí
AV75	BOG	LIM	21,92	0,83	Sí
AV24	LIM	BOG	17,50	20,67	Sí
AV10	BOG	MAD	21,58	14,25	Sí
AV11	MAD	BOG	16,42	20,00	Sí
AV9302	BOG	MDE	6,00	6,97	Sí
AV9301	MDE	BOG	6,00	6,83	Sí

AV9322	BOG	MDE	6,50	7,47	Sí
AV9303	MDE	BOG	6,50	7,33	Sí
AV9335	MDE	BOG	6,75	7,58	Sí
AV9304	BOG	MDE	7,08	8,05	No
AV8431	MDE	BOG	7,00	7,83	Sí
AV9324	BOG	MDE	8,00	8,97	Sí
AV9305	MDE	BOG	7,50	8,33	No
AV9310	BOG	MDE	9,00	9,97	Sí
AV9323	MDE	BOG	8,00	8,83	Sí
AV9338	BOG	MDE	10,00	10,95	Sí
AV9307	MDE	BOG	8,58	9,42	Sí
AV8464	BOG	MDE	10,00	11,00	Sí
AV9319	MDE	BOG	9,75	10,67	Sí
AV9024	BOG	MDE	11,00	11,95	Sí
AV9311	MDE	BOG	10,50	11,33	No
AV9312	BOG	MDE	11,08	12,08	Sí
AV8465	MDE	BOG	11,50	12,33	Sí
AV8424	BOG	MDE	12,00	12,95	Sí
AV9339	MDE	BOG	11,50	12,33	Sí
AV9308	BOG	MDE	13,00	13,95	Sí
AV9313	MDE	BOG	12,58	13,42	Sí
AV9316	BOG	MDE	14,00	14,95	Sí
AV9025	MDE	BOG	12,58	13,42	Sí
AV9318	BOG	MDE	14,50	15,62	Sí
AV8425	MDE	BOG	13,50	14,33	Sí
AV9348	BOG	MDE	15,00	15,95	No
AV9309	MDE	BOG	14,50	15,33	No
AV8436	BOG	MDE	15,00	16,00	No
AV9317	MDE	BOG	15,50	16,33	Sí
AV9314	BOG	MDE	15,92	16,87	Sí
AV8437	MDE	BOG	16,50	17,33	No
AV9026	BOG	MDE	16,00	16,95	Sí
AV9349	MDE	BOG	16,50	17,33	Sí
AV8438	BOG	MDE	16,00	17,00	Sí
AV8439	MDE	BOG	17,50	18,33	Sí
AV9016	BOG	MDE	16,25	17,08	No
AV9315	MDE	BOG	17,50	18,33	Sí
AV9330	BOG	MDE	16,50	17,45	Sí
AV9027	MDE	BOG	17,50	18,33	Sí
AV8462	BOG	MDE	17,00	18,00	Sí

AV9331	MDE	BOG	18,08	18,92	Sí
AV9320	BOG	MDE	17,03	18,00	Sí
AV9017	MDE	BOG	18,25	19,08	Sí
AV8463	MDE	BOG	18,50	19,33	Sí
AV8434	BOG	MDE	17,50	18,50	Sí
AV9321	MDE	BOG	18,50	19,33	Sí
AV9336	BOG	MDE	18,00	18,97	Sí
AV9028	MDE	BOG	18,58	19,42	Sí
AV8426	BOG	MDE	18,00	19,00	No
AV9333	MDE	BOG	19,00	19,83	Sí
AV9326	BOG	MDE	19,03	20,00	Sí
AV8435	MDE	BOG	19,17	20,00	Sí
AV9328	BOG	MDE	20,00	20,87	Sí
AV8427	MDE	BOG	19,50	20,33	Sí
AV9337	MDE	BOG	19,50	20,33	Sí
AV8428	BOG	MDE	20,00	21,00	Sí
AV9327	MDE	BOG	20,50	21,33	Sí
AV9334	BOG	MDE	20,25	21,12	No
AV8429	MDE	BOG	21,50	22,33	Sí
AV9332	BOG	MDE	20,50	21,37	Sí
AV9329	MDE	BOG	21,50	22,33	Sí
AV9300	BOG	MDE	21,50	22,37	Sí
AV8430	BOG	MDE	22,00	23,00	No
AV72	BOG	MEX	9,67	14,33	Sí
AV23	MEX	BOG	1,58	6,00	Sí
AV22	BOG	MEX	14,50	19,17	Sí
AV45	MEX	BOG	8,50	12,92	Sí
AV44	BOG	MEX	22,67	3,33	Sí
AV73	MEX	BOG	15,75	20,17	Sí
AV6	BOG	MIA	10,33	15,00	Sí
AV9	MIA	BOG	7,37	10,00	Sí
AV7	MIA	BOG	17,00	19,50	Sí
AV8	BOG	MIA	18,25	22,92	Sí
AV8506	BOG	MTR	8,92	10,30	No
AV8507	MTR	BOG	10,80	11,97	No
AV9440	BOG	MTR	12,25	13,63	No
AV9441	MTR	BOG	14,08	15,33	No
AV8508	BOG	MTR	15,33	16,73	No
AV8509	MTR	BOG	17,25	18,50	No
AV9926	BOG	MZL	5,83	6,77	No

AV9927	MZL	BOG	7,03	7,82	No
AV9920	BOG	MZL	7,00	8,00	No
AV9921	MZL	BOG	8,25	9,03	No
AV9922	BOG	MZL	9,50	10,52	No
AV9923	MZL	BOG	10,92	11,70	No
AV9928	BOG	MZL	11,08	12,02	No
AV9929	MZL	BOG	12,27	13,07	No
AV9924	BOG	MZL	13,40	14,40	No
AV9925	MZL	BOG	14,67	15,47	No
AV9932	BOG	MZL	15,67	16,70	No
AV9933	MZL	BOG	17,03	17,87	No
AV9930	BOG	MZL	16,25	17,33	No
AV9931	MZL	BOG	17,67	18,50	No
AV9285	BOG	NVA	6,08	7,18	No
AV9286	NVA	BOG	7,52	8,48	No
AV9029	BOG	NVA	7,95	9,05	No
AV9030	NVA	BOG	9,38	10,35	No
AV9291	BOG	NVA	9,00	10,05	No
AV9292	NVA	BOG	10,50	11,47	No
AV9289	BOG	NVA	13,58	14,55	No
AV9290	NVA	BOG	14,92	15,92	No
AV9031	BOG	NVA	17,42	18,45	No
AV9032	NVA	BOG	18,83	19,83	No
AV9287	BOG	NVA	19,08	20,12	No
AV9288	NVA	BOG	20,50	21,50	No
AV8503	BOG	PEI	6,08	7,17	No
AV8504	PEI	BOG	6,00	6,83	No
AV9905	BOG	PEI	7,83	8,75	No
AV8502	PEI	BOG	7,58	8,42	No
AV8511	BOG	PEI	9,42	10,42	No
AV9906	PEI	BOG	9,00	9,87	No
AV8515	BOG	PEI	12,25	13,25	No
AV8512	PEI	BOG	10,92	11,75	No
AV8513	BOG	PEI	14,75	15,75	No
AV8516	PEI	BOG	13,67	14,50	No
AV8517	BOG	PEI	15,75	16,75	No
AV8514	PEI	BOG	16,17	17,00	No
AV9911	BOG	PEI	18,08	19,12	No
AV8518	PEI	BOG	17,25	18,08	No
AV9903	BOG	PEI	19,33	20,37	No
AV9918	PEI	BOG	18,67	19,50	No

AV9917	BOG	PEI	20,83	21,67	No
AV9912	PEI	BOG	19,58	20,55	No
AV8505	BOG	PEI	22,00	23,08	No
AV9904	PEI	BOG	20,75	21,72	No
AV9241	BOG	PPN	8,17	9,50	No
AV9242	PPN	BOG	9,92	11,25	No
AV9243	BOG	PPN	15,92	17,25	No
AV9244	PPN	BOG	17,67	19,00	No
AV8583	BOG	PSO	8,25	9,67	No
AV8584	PSO	BOG	10,08	11,33	No
AV9401	BOG	PSO	11,92	13,50	No
AV9402	PSO	BOG	13,75	15,33	No
AV8597	BOG	PSO	14,88	16,25	No
AV8598	PSO	BOG	16,75	18,00	No
AV58	BOG	PTY	7,50	9,08	Sí
AV59	PTY	BOG	10,00	11,45	Sí
AV60	BOG	PTY	14,42	15,92	Sí
AV61	PTY	BOG	17,92	19,33	Sí
AV82	BOG	PUJ	20,75	0,42	Sí
AV83	PUJ	BOG	1,32	3,10	Sí
AV8590	BOG	RCH	10,92	12,58	No
AV8591	RCH	BOG	13,00	14,50	No
AV9412	BOG	RCH	10,92	12,58	No
AV9413	RCH	BOG	13,00	14,50	No
AV94	BOG	SAL	11,92	14,00	Sí
AV95	SAL	BOG	14,92	19,00	Sí
AV97	BOG	SCL	21,83	5,75	Sí
AV98	SCL	BOG	8,42	12,42	Sí
AV250	BOG	SDQ	14,00	17,62	Sí
AV251	SDQ	BOG	18,50	20,17	Sí
AV8484	BOG	SMR	7,25	8,82	No
AV8485	SMR	BOG	9,25	10,67	No
AV9774	BOG	SMR	11,25	12,75	No
AV9775	SMR	BOG	13,25	14,68	No
AV9766	BOG	SMR	15,27	16,77	No
AV9767	SMR	BOG	17,35	18,78	No
AV9768	BOG	SMR	19,42	20,92	No
AV9769	SMR	BOG	21,42	22,85	No
AV6371	BOG	UIO	8,75	10,30	Sí
AV6378	UIO	BOG	6,33	7,75	Sí
AV6373	BOG	UIO	14,17	15,58	Sí

AV6372	UIO	BOG	11,38	12,83	Sí
AV6375	BOG	UIO	18,75	20,17	Sí
AV6374	UIO	BOG	16,33	17,78	Sí
AV6377	BOG	UIO	22,00	23,55	Sí
AV6376	UIO	BOG	18,83	20,35	Sí
AV8596	BOG	VUP	8,83	10,33	No
AV8595	VUP	BOG	10,83	12,17	No
AV9410	BOG	VUP	12,58	14,00	No
AV9411	VUP	BOG	14,42	15,75	No
AV8592	BOG	VUP	17,75	19,22	No
AV8593	VUP	BOG	19,67	21,00	No
AV14	CLO	MAD	17,92	11,00	Sí
AV15	MAD	CLO	13,50	17,33	Sí
AV9433	CLO	MDE	6,00	7,05	No
AV9434	MDE	CLO	6,08	7,25	No
AV8525	CLO	MDE	6,92	7,75	No
AV8524	MDE	CLO	8,17	9,17	No
AV9435	CLO	MDE	12,92	13,97	No
AV9033	MDE	CLO	14,35	15,28	No
AV9431	CLO	MDE	18,92	19,97	No
AV9430	MDE	CLO	15,17	16,23	No
AV9034	CLO	MDE	21,92	22,75	No
AV9432	MDE	CLO	20,33	21,43	No
AV38	CLO	MIA	10,25	14,83	Sí
AV39	MIA	CLO	17,35	20,00	Sí

AV9403	CLO	PSO	10,50	11,45	No
AV9404	PSO	CLO	11,70	12,57	No
AV9405	CLO	PSO	16,50	17,45	No
AV9406	PSO	CLO	17,75	18,60	No
AV9419	CLO	TCO	7,63	8,80	No
AV9420	TCO	CLO	9,17	10,25	No
AV9761	CTG	MDE	9,00	10,42	No
AV9760	MDE	CTG	7,33	8,70	No
AV9511	CTG	MDE	16,08	17,58	No
AV9510	MDE	CTG	14,50	15,83	No
AV9771	CTG	MDE	20,08	21,63	No
AV9770	MDE	CTG	18,50	19,83	No
AV34	CTG	MIA	9,17	12,92	Sí
AV35	MIA	CTG	14,17	15,77	Sí
AV9035	CTG	PEI	16,83	18,08	No
AV9036	PEI	CTG	7,00	8,17	No
AV43	JFK	MDE	9,00	13,60	Sí
AV42	MDE	JFK	23,92	6,17	Sí
AV17	MAD	MDE	13,17	17,00	Sí
AV16	MDE	MAD	18,33	11,17	Sí
AV30	MDE	MIA	10,00	14,25	Sí
AV31	MIA	MDE	15,75	18,00	Sí
AV9360	MDE	SMR	10,67	12,22	No
AV9361	SMR	MDE	12,50	14,10	No

Distancia y demanda entre las ciudades consideradas en el itinerario base

Origen	Destino	Distancia (mi)	Demanda promedio por vuelo
ADZ	BOG	750	140
BOG	ADZ	750	140
ADZ	CTG	446	140
CTG	ADZ	446	140
AUA	BOG	606	130
BOG	AUA	606	130
AXM	BOG	114	135
BOG	AXM	114	135
BAQ	BOG	428	140
BOG	BAQ	428	140

BAQ	MDE	328	45
MDE	BAQ	328	45
BAQ	MIA	1087	140
MIA	BAQ	1087	140
BCN	BOG	5292	250
BOG	BCN	5292	250
BGA	BOG	180	140
BOG	BGA	180	140
BOG	CCS	636	130
CCS	BOG	636	130
BOG	CLO	175	140
CLO	BOG	175	140
BOG	CTG	406	140

CTG	BOG	406	140
BOG	CUC	249	130
CUC	BOG	249	130
BOG	CUR	625	120
CUR	BOG	625	120
BOG	EJA	162	35
EJA	BOG	162	35
BOG	EZE	2903	250
EZE	BOG	2903	250
BOG	FLL	1523	140
FLL	BOG	1523	140
BOG	GRU	2689	200
GRU	BOG	2689	200
BOG	GYE	617	130
GYE	BOG	617	130
BOG	IAD	2366	140
IAD	BOG	2366	140
BOG	IBE	72	35
IBE	BOG	72	35
BOG	JFK	2474	250
JFK	BOG	2474	250
BOG	LAX	3479	250
LAX	BOG	3479	250
BOG	LIM	1168	250
LIM	BOG	1168	250
BOG	MAD	4992	250
MAD	BOG	4992	250
BOG	MDE	134	150
MDE	BOG	134	150
BOG	MEX	1963	190
MEX	BOG	1963	190
BOG	MIA	1507	250
MIA	BOG	1507	250
BOG	MTR	306	150
MTR	BOG	306	150
BOG	MZL	95	35
MZL	BOG	95	35
BOG	NVA	145	35
NVA	BOG	145	35

BOG	PEI	111	140
PEI	BOG	111	140
BOG	PPN	231	35
PPN	BOG	231	35
BOG	PSO	315	115
PSO	BOG	315	115
BOG	PTY	469	230
PTY	BOG	469	230
BOG	PUJ	1030	175
PUJ	BOG	1030	175
BOG	RCH	476	120
RCH	BOG	476	120
BOG	SAL	1182	130
SAL	BOG	1182	130
BOG	SCL	2630	220
SCL	BOG	2630	220
BOG	SDQ	991	165
SDQ	BOG	991	165
BOG	SMR	442	145
SMR	BOG	442	145
BOG	UIO	442	130
UIO	BOG	442	130
BOG	VUP	399	130
VUP	BOG	399	130
CLO	MAD	5161	250
MAD	CLO	5161	250
CLO	MDE	192	140
MDE	CLO	192	140
CLO	MIA	1552	140
MIA	CLO	1552	140
CLO	PSO	161	35
PSO	CLO	161	35
CLO	TCO	203	35
TCO	CLO	203	35
CTG	MDE	295	35
MDE	CTG	295	35
CTG	MIA	1102	120
MIA	CTG	1102	120
CTG	PEI	388	100

PEI	CTG	388	100
JFK	MDE	2376	125
MDE	JFK	2376	125
MAD	MDE	4993	250
MDE	MAD	4993	250

MDE	MIA	1388	150
MIA	MDE	1388	150
MDE	SMR	351	35
SMR	MDE	351	35

C. Anexo: Resultados del caso de aplicación

Vuelo	Comparación de horas de salida y de tipos de avión asignados						
	Itinerario Avianca		Modelo propuesto sin clusterización		Modelo propuesto con clusterización		
	Hora de salida	Tipo de avión	Hora de salida	Tipo de avión	Hora de salida	Clúster de aviones	Desclusterizado
AV9779	14,08	A320	13,58	A320	13,58	pSencillo	A320
AV9778	11,50	A320	11,50	A320	12,00	pSencillo	A320
AV9777	16,83	A320	16,83	A320	16,33	pSencillo	A320
AV9776	14,17	A320	14,67	A320	14,67	pSencillo	A320
AV9781	22,00	A318	22,50	A320	22,00	pSencillo	A320
AV9780	19,25	A318	19,75	A320	19,75	pSencillo	A320
AV9001	12,08	A318	12,08	A320	11,58	pSencillo	A320
AV9002	10,17	A318	9,67	A320	10,67	pSencillo	A320
AV9004	13,33	A318	13,83	A320	12,83	pSencillo	A320
AV9003	11,42	A318	11,92	A320	11,92	pSencillo	A320
AV91	18,67	A318	18,17	A319	18,17	pSencillo	A319
AV90	14,33	A320	14,83	A319	14,83	pSencillo	A319
AV9940	7,35	F50	6,85	A318	6,85	pSencillo	A319
AV9939	6,00	F50	6,50	A320	6,00	pSencillo	A319
AV9938	9,83	F50	9,33	A320	9,33	pSencillo	A319
AV9937	8,58	F50	9,08	A319	9,08	pSencillo	A319
AV9942	12,33	F50	11,83	A319	11,83	pSencillo	A319
AV9941	11,08	F50	11,58	A319	11,58	pSencillo	A320
AV9944	16,80	F50	16,30	A319	16,30	pSencillo	A320
AV9943	15,42	F50	15,92	A319	15,92	pSencillo	A319
AV9936	19,50	F50	19,00	A319	19,00	pSencillo	A319
AV9935	18,25	F50	18,75	A318	18,75	pSencillo	A319
AV9521	6,08	A318	5,58	A320	5,58	pSencillo	A320
AV8522	5,78	A318	5,28	A320	5,28	pSencillo	A320
AV8533	7,00	A318	6,50	A320	6,50	pSencillo	A320
AV9536	6,25	A320	5,75	A320	5,75	pSencillo	A320
AV9005	7,50	A318	Vuelo no elegido	Vuelo no elegido	7,00	pSencillo	A320
AV9522	7,50	A320	8,00	A320	8,00	pSencillo	A320

AV8523	7,83	A318	7,33	A320	7,33	pSencillo	A320
AV9534	8,50	A320	9,00	A320	9,00	pSencillo	A320
AV9537	8,33	A320	7,83	A320	7,83	pSencillo	A320
AV8520	9,75	A318	10,25	A320	10,25	pSencillo	A320
AV9006	8,33	A318	7,83	A320	7,83	pSencillo	A320
AV9007	10,42	A318	10,92	A320	10,92	pSencillo	A320
AV9533	10,67	A320	10,17	A320	10,17	pSencillo	A318
AV9538	11,92	A320	12,42	A320	12,42	pSencillo	A318
AV8521	11,75	A318	11,75	A320	11,25	pSencillo	A320
AV8528	15,25	A318	15,75	A320	15,75	pSencillo	A320
AV9008	12,42	A318	12,42	A320	11,92	pSencillo	A320
AV8532	17,42	A318	17,92	A320	17,92	pSencillo	A320
AV9539	13,83	A320	13,83	A320	13,33	pSencillo	A320
AV9530	19,67	A320	20,17	A320	20,17	pSencillo	A320
AV8529	17,33	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9532	20,83	A318	21,33	A320	21,33	pSencillo	A320
AV9535	17,95	A320	17,95	A320	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV8530	21,92	A318	22,42	A320	22,42	pSencillo	A320
AV8531	19,42	A318	18,92	A320	18,92	pSencillo	A320
AV9531	21,58	A320	21,08	A320	21,08	pSencillo	A320
AV9483	7,83	F50	7,83	F50	7,83	pSencillo	F50
AV9482	6,00	F50	5,50	F50	5,50	pSencillo	F50
AV9369	18,17	F50	17,67	F50	17,67	pSencillo	A318
AV9368	16,42	F50	15,92	F50	16,92	pSencillo	A318
AV2	10,17	A320	10,67	A320	10,17	pSencillo	A320
AV3	15,17	A320	15,67	A320	15,17	pSencillo	A320
AV19	16,67	A330	16,17	A330	16,17	pDoble	A330
AV18	21,42	A330	21,92	A330	21,92	pDoble	A330
AV9471	6,00	A318	5,50	A320	5,50	pSencillo	A320
AV8572	6,17	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV8447	7,75	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9009	7,42	A319	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9010	9,00	A319	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9460	7,42	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9461	9,00	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9472	9,17	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9473	10,73	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9468	10,50	F50	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9469	11,92	F50	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido

AV8578	12,50	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV8579	13,92	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9464	14,92	A320	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9465	16,50	A320	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV8564	15,08	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV8565	16,58	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9462	17,17	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9463	18,58	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV8566	18,42	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV8567	19,83	A318	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9466	19,50	A320	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9467	21,00	A320	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9470	21,92	A318	22,42	A320	22,42	pSencillo	A320
AV66	8,50	A319	9,00	A319	9,00	pSencillo	A320
AV79	6,25	A319	5,75	A320	5,75	pSencillo	A319
AV9011	8,50	A318	8,50	A320	9,00	pSencillo	A319
AV69	11,75	A318	11,25	A320	11,25	pSencillo	A319
AV80	14,42	A319	14,92	A319	14,92	pSencillo	A320
AV9012	11,92	A319	11,42	A319	11,42	pSencillo	A320
AV78	21,67	A319	22,17	A320	22,17	pSencillo	A319
AV81	18,67	A319	18,17	A319	18,17	pSencillo	A320
AV9201	6,00	A320	6,50	A320	5,50	pSencillo	A320
AV9200	6,00	A320	5,50	A320	5,50	pSencillo	A320
AV9217	6,25	A319	6,25	A320	5,75	pSencillo	A320
AV9236	6,50	A320	6,00	A320	6,00	pSencillo	A320
AV8401	6,50	A318	6,50	A320	6,00	pSencillo	A320
AV9220	7,00	A320	6,50	A320	6,50	pSencillo	A320
AV9203	6,92	A320	7,42	A320	7,42	pSencillo	A320
AV9202	7,58	A320	7,08	A319	7,08	pSencillo	A320
AV9231	8,00	A320	8,50	A320	8,50	pSencillo	A320
AV8412	7,58	A318	7,08	A320	7,08	pSencillo	A320
AV9207	8,92	A320	8,92	A320	9,42	pSencillo	A320
AV8400	8,08	A318	8,08	A320	7,58	pSencillo	A318
AV9209	10,00	A320	10,50	A320	10,50	pSencillo	A320
AV9218	8,33	A319	7,83	A320	7,83	pSencillo	A320
AV9204	8,50	A320	8,00	A320	8,00	pSencillo	A320
AV9211	11,00	A320	11,50	A320	11,50	pSencillo	A320
AV8418	9,58	A318	9,08	A320	9,08	pSencillo	A320
AV8409	12,00	A318	12,50	A320	12,50	pSencillo	A320
AV9208	10,58	A320	10,08	A320	10,08	pSencillo	A320

AV9221	13,00	A320	13,50	A320	13,50	pSencillo	A320
AV9205	14,00	A320	14,50	A320	14,50	pSencillo	A318
AV9210	11,67	A320	Vuelo no elegido	Vuelo no elegido	11,17	pSencillo	A320
AV9212	12,58	A320	12,08	A320	12,08	pSencillo	A318
AV8443	14,00	A318	14,50	A320	14,50	pSencillo	A320
AV8420	13,58	A318	13,08	A318	13,08	pSencillo	A320
AV8421	15,00	A318	15,50	A320	15,50	pSencillo	A320
AV9222	14,58	A320	14,08	A320	14,08	pSencillo	A320
AV9013	15,58	A330	16,08	A320	16,08	pSencillo	A320
AV9206	15,58	A320	15,08	A320	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9213	15,92	A320	16,42	A320	16,42	pSencillo	A320
AV9215	17,00	A320	17,00	A319	17,50	pSencillo	A319
AV8442	15,58	A318	15,08	A320	15,08	pSencillo	A320
AV9227	18,00	A319	18,50	A320	18,50	pSencillo	A318
AV9226	16,12	A319	15,62	A320	15,62	pSencillo	A320
AV8405	18,50	A318	19,00	A318	19,00	pSencillo	A320
AV8422	16,67	A318	16,17	A320	16,17	pSencillo	A320
AV9233	19,08	A318	19,58	A320	19,58	pSencillo	A320
AV9214	17,58	A320	17,58	A320	17,08	pSencillo	A320
AV9223	20,00	A320	20,50	A320	20,50	pSencillo	A320
AV9014	18,58	A330	18,08	A320	18,08	pSencillo	A319
AV9219	20,50	A320	21,00	A320	21,00	pSencillo	A320
AV9216	19,00	A320	18,50	A319	18,50	pSencillo	A320
AV9015	21,17	A320	21,67	A320	21,67	pSencillo	A320
AV8406	20,08	A318	19,58	A320	19,58	pSencillo	A320
AV8417	22,00	A318	22,00	A320	22,50	pSencillo	A320
AV9234	20,67	A318	20,17	A318	20,17	pSencillo	A320
AV9224	21,58	A320	21,08	A320	21,58	pSencillo	A320
AV9540	7,42	A320	7,92	A320	7,92	pSencillo	A320
AV9755	6,00	A320	5,50	A320	5,50	pSencillo	A320
AV9786	8,17	A318	8,67	A320	8,67	pSencillo	A320
AV9541	9,33	A320	8,83	A320	8,83	pSencillo	A320
AV9542	8,75	A320	9,25	A320	9,25	pSencillo	A320
AV9543	10,67	A320	10,17	A320	10,17	pSencillo	A320
AV8490	9,42	A318	9,92	A320	9,92	pSencillo	A320
AV9549	11,42	A320	10,92	A320	10,92	pSencillo	A320
AV9548	9,50	A320	10,00	A320	10,00	pSencillo	A320
AV9018	12,17	A320	11,67	A320	11,67	pSencillo	A320
AV9019	10,25	A320	10,75	A320	10,75	pSencillo	A320

AV9759	12,25	A320	12,25	A320	12,75	pSencillo	A320
AV9758	10,33	A320	10,83	A320	10,83	pSencillo	A320
AV9753	14,00	A320	13,50	A320	13,50	pSencillo	A320
AV9752	12,08	A320	12,08	A320	12,58	pSencillo	A320
AV9787	14,00	A318	13,50	A320	13,50	pSencillo	A320
AV9788	13,00	A318	13,00	A320	13,50	pSencillo	A320
AV9789	14,92	A318	14,42	A320	14,42	pSencillo	A320
AV9546	13,67	A320	14,17	A320	14,17	pSencillo	A320
AV8491	15,25	A318	14,75	A320	14,75	pSencillo	A320
AV9544	15,83	A320	15,83	A318	16,33	pSencillo	A320
AV9547	15,58	A320	16,08	A320	15,08	pSencillo	A320
AV8548	18,08	A318	18,58	A320	18,58	pSencillo	A320
AV9545	17,75	A320	Vuelo no elegido	Vuelo no elegido	17,25	pSencillo	A320
AV8550	18,92	A318	19,42	A320	19,42	pSencillo	A320
AV8549	20,00	A318	20,00	A320	19,50	pSencillo	A320
AV8552	20,42	A318	20,92	A320	20,92	pSencillo	A320
AV8551	20,83	A318	20,33	A320	20,33	pSencillo	A320
AV9754	21,92	A320	22,42	A320	22,42	pSencillo	A320
AV8553	22,33	A318	22,83	A320	22,83	pSencillo	A320
AV9450	6,08	A320	5,58	A318	6,08	pSencillo	A318
AV9451	7,75	A320	7,25	A319	7,25	pSencillo	A318
AV9448	9,17	A320	9,67	A319	9,67	pSencillo	A319
AV9449	10,92	A320	10,42	A318	10,42	pSencillo	A319
AV9020	12,25	F50	12,75	A319	12,75	pSencillo	A319
AV9021	14,08	F50	13,58	A319	13,58	pSencillo	A319
AV9456	13,92	F50	14,42	A319	14,42	pSencillo	A319
AV9457	15,75	F50	15,25	A319	15,25	pSencillo	A319
AV9452	17,50	A320	18,00	A318	18,00	pSencillo	A320
AV9453	19,17	A320	18,67	A319	18,67	pSencillo	A319
AV9458	19,00	A320	19,50	A319	19,50	pSencillo	A319
AV9459	20,67	A320	20,17	A318	20,17	pSencillo	A320
AV92	14,67	A318	15,17	A319	15,17	pSencillo	A319
AV93	19,00	A318	18,50	A319	18,50	pSencillo	A319
AV8586	6,08	A318	6,58	F50	6,58	pSencillo	F50
AV8587	7,68	A318	7,18	F50	7,18	pSencillo	F50
AV9494	11,58	F50	12,08	F50	12,08	pSencillo	F50
AV9495	13,08	F50	12,58	F50	12,58	pSencillo	F50
AV9492	15,92	F50	16,42	F50	16,42	pSencillo	F50
AV9493	17,42	F50	16,92	F50	16,92	pSencillo	F50
AV8580	16,17	A318	16,17	F50	16,67	pSencillo	F50

AV8581	17,67	A318	17,17	F50	17,67	pSencillo	F50
AV87	22,17	A330	22,67	A330	22,67	pDoble	A330
AV88	7,75	A330	7,25	A330	7,25	pDoble	A330
AV36	9,73	A320	10,23	A320	10,23	pSencillo	A320
AV37	17,33	A320	16,83	A320	16,83	pSencillo	A320
AV85	21,60	A330	22,10	A320	22,10	pDoble	A330
AV86	8,58	A330	8,08	A320	8,08	pDoble	A330
AV113	14,67	A318	15,17	A319	15,17	pSencillo	A320
AV280	6,17	A320	5,67	A320	5,67	pSencillo	A319
AV281	21,75	A320	22,25	A320	22,25	pSencillo	A319
AV112	18,00	A318	17,50	A319	17,50	pSencillo	A320
AV286	9,35	A319	9,85	A320	9,85	pSencillo	A320
AV287	16,60	A319	16,10	A320	16,10	pSencillo	A320
AV9281	5,83	F50	5,33	F50	6,33	pSencillo	F50
AV9282	6,92	F50	6,42	F50	6,42	pSencillo	F50
AV9279	13,32	F50	13,82	A318	13,82	pSencillo	F50
AV9280	14,42	F50	13,92	A319	13,92	pSencillo	F50
AV9022	15,33	F50	15,83	A318	15,83	pSencillo	F50
AV9023	16,42	F50	15,92	A318	15,92	pSencillo	F50
AV9283	17,42	F50	17,92	A319	17,92	pSencillo	F50
AV9284	18,50	F50	18,00	A318	18,00	pSencillo	F50
AV284	14,72	A319	15,22	B767	15,22	pDoble	B767
AV285	0,92	A319	0,42	B767	0,42	pDoble	B767
AV20	22,92	B767	23,42	B767	23,42	pDoble	B767
AV21	8,00	B767	7,50	B767	7,50	pDoble	B767
AV48	1,58	B767	2,08	A330	2,08	pDoble	A330
AV49	11,50	B767	11,00	A330	11,00	pDoble	A330
AV25	13,67	A320	13,67	A330	13,67	pDoble	A330
AV74	5,83	B767	5,33	B767	5,33	pDoble	B767
AV75	21,92	B767	21,92	B767	22,42	pDoble	B767
AV24	17,50	A320	17,00	A330	17,00	pDoble	A330
AV10	21,58	A330	22,08	A330	22,08	pDoble	A330
AV11	16,42	A330	15,92	A330	15,92	pDoble	A330
AV9302	6,00	A320	5,50	A320	6,00	pSencillo	A320
AV9301	6,00	A320	5,50	A320	5,50	pSencillo	A320
AV9322	6,50	A320	7,00	A320	6,00	pSencillo	A320
AV9303	6,50	A320	6,00	A320	6,00	pSencillo	A320
AV9335	6,75	A320	6,25	A320	6,25	pSencillo	A320
AV9304	7,08	A320	7,58	A320	7,58	pSencillo	A320

AV8431	7,00	A318	6,50	A320	6,50	pSencillo	A320
AV9324	8,00	A320	8,50	A320	8,50	pSencillo	A320
AV9305	7,50	A320	7,00	A320	7,00	pSencillo	A320
AV9310	9,00	A320	9,50	A320	9,50	pSencillo	A320
AV9323	8,00	A320	7,50	A320	7,50	pSencillo	A320
AV9338	10,00	A320	10,00	A320	10,50	pSencillo	A320
AV9307	8,58	A320	8,08	A320	8,08	pSencillo	A320
AV8464	10,00	A318	10,50	A320	10,50	pSencillo	A320
AV9319	9,75	F50	9,25	A320	9,25	pSencillo	A320
AV9024	11,00	A320	11,50	A320	11,50	pSencillo	A319
AV9311	10,50	A320	10,00	A320	11,00	pSencillo	A320
AV9312	11,08	A318	11,58	A320	11,58	pSencillo	A320
AV8465	11,50	A318	11,00	A320	11,00	pSencillo	A319
AV8424	12,00	A318	12,50	A320	12,50	pSencillo	A320
AV9339	11,50	A320	11,00	A318	11,00	pSencillo	A318
AV9308	13,00	A320	13,50	A320	13,50	pSencillo	A320
AV9313	12,58	A318	12,08	A320	12,08	pSencillo	A320
AV9316	14,00	A320	14,50	A320	14,50	pSencillo	A320
AV9025	12,58	A320	12,58	A320	12,08	pSencillo	A320
AV9318	14,50	F50	15,00	A320	14,50	pSencillo	A320
AV8425	13,50	A318	13,00	A320	13,00	pSencillo	A320
AV9348	15,00	A320	15,50	A319	15,50	pDoble	B767
AV9309	14,50	A320	14,50	A320	14,00	pSencillo	A320
AV8436	15,00	A318	15,00	A320	15,50	pSencillo	A320
AV9317	15,50	A320	15,00	A320	15,00	pSencillo	A320
AV9314	15,92	A320	16,42	A318	16,42	pSencillo	A320
AV8437	16,50	A318	16,00	A320	16,50	pSencillo	A320
AV9026	16,00	A320	16,50	A320	16,50	pSencillo	A320
AV9349	16,50	A320	16,00	A320	16,00	pSencillo	A320
AV8438	16,00	A318	16,50	A320	16,50	pSencillo	A320
AV8439	17,50	A318	17,00	A320	17,00	pSencillo	A320
AV9016	16,25	A330	16,75	A320	16,75	pSencillo	A320
AV9315	17,50	A320	17,00	A319	17,00	pSencillo	A319
AV9330	16,50	A320	17,00	A320	17,00	pSencillo	A320
AV9027	17,50	A320	18,00	A318	17,00	pDoble	B767
AV8462	17,00	A318	16,50	A320	17,50	pSencillo	A320
AV9331	18,08	A320	18,08	A320	18,08	pSencillo	A320
AV9320	17,03	A320	17,53	A320	17,53	pSencillo	A320
AV9017	18,25	A330	18,75	A320	18,25	pSencillo	A320
AV8463	18,50	A318	18,00	A320	18,00	pSencillo	A320

AV8434	17,50	A318	17,00	A320	17,50	pSencillo	A320
AV9321	18,50	A320	18,00	A320	18,50	pSencillo	A320
AV9336	18,00	A320	18,50	A320	18,50	pSencillo	A320
AV9028	18,58	A320	18,08	A320	18,08	pSencillo	A320
AV8426	18,00	A318	18,00	A320	18,50	pSencillo	A320
AV9333	19,00	A320	18,50	A320	19,00	pSencillo	A320
AV9326	19,03	A320	19,53	A320	19,53	pSencillo	A320
AV8435	19,17	A318	19,67	A320	19,17	pSencillo	A320
AV9328	20,00	A320	20,50	A320	20,50	pSencillo	A320
AV8427	19,50	A318	19,50	A320	19,00	pSencillo	A320
AV9337	19,50	A320	19,00	A320	19,00	pSencillo	A320
AV8428	20,00	A318	20,50	A320	20,50	pSencillo	A320
AV9327	20,50	A320	20,00	A320	20,00	pSencillo	A320
AV9334	20,25	A320	20,75	A320	20,75	pSencillo	A319
AV8429	21,50	A318	22,00	A320	21,00	pSencillo	A320
AV9332	20,50	A320	21,00	A318	21,00	pSencillo	A320
AV9329	21,50	A320	21,00	A320	21,00	pSencillo	A320
AV9300	21,50	A320	22,00	A320	22,00	pSencillo	A320
AV8430	22,00	A318	22,50	A319	22,50	pSencillo	A320
AV72	9,67	A320	10,17	A320	10,17	pDoble	B767
AV23	1,58	A320	1,08	A320	1,08	pSencillo	A320
AV22	14,50	A320	15,00	A320	15,00	pSencillo	A320
AV45	8,50	A319	8,00	A320	8,00	pSencillo	A320
AV44	22,67	A319	23,17	A320	23,17	pSencillo	A320
AV73	15,75	A320	15,25	A320	15,25	pDoble	B767
AV6	10,33	B767	10,83	B767	10,83	pDoble	B767
AV9	7,37	A320	6,87	B767	6,87	pDoble	B767
AV7	17,00	B767	16,50	B767	16,50	pDoble	B767
AV8	18,25	A320	18,75	B767	18,75	pDoble	B767
AV8506	8,92	A318	9,42	A320	9,42	pSencillo	A320
AV8507	10,80	A318	10,30	A320	10,30	pSencillo	A320
AV9440	12,25	A318	12,75	A320	12,75	pSencillo	A320
AV9441	14,08	A318	13,58	A320	13,58	pSencillo	A320
AV8508	15,33	A318	15,83	A320	15,83	pSencillo	A320
AV8509	17,25	A318	16,75	A320	16,75	pSencillo	A320
AV9926	5,83	F50	5,83	F50	5,83	pSencillo	F50
AV9927	7,03	F50	6,53	F50	6,53	pSencillo	F50
AV9920	7,00	F50	7,50	F50	7,50	pSencillo	F50
AV9921	8,25	F50	7,75	F50	7,75	pSencillo	F50

AV9922	9,50	F50	10,00	F50	10,00	pSencillo	F50
AV9923	10,92	F50	10,42	F50	10,42	pSencillo	F50
AV9928	11,08	F50	11,58	F50	11,58	pSencillo	F50
AV9929	12,27	F50	11,77	F50	11,77	pSencillo	F50
AV9924	13,40	F50	13,90	F50	13,40	pSencillo	F50
AV9925	14,67	F50	14,17	F50	14,17	pSencillo	F50
AV9932	15,67	F50	16,17	F50	16,17	pSencillo	F50
AV9933	17,03	F50	16,53	F50	16,53	pSencillo	F50
AV9930	16,25	F50	16,75	A318	16,75	pSencillo	A319
AV9931	17,67	F50	17,17	A318	17,17	pSencillo	A319
AV9285	6,08	F50	6,58	F50	6,58	pSencillo	F50
AV9286	7,52	F50	7,02	F50	7,02	pSencillo	F50
AV9029	7,95	F50	8,45	F50	8,45	pSencillo	F50
AV9030	9,38	F50	8,88	F50	8,88	pSencillo	F50
AV9291	9,00	F50	9,50	F50	9,50	pSencillo	F50
AV9292	10,50	F50	10,00	F50	10,00	pSencillo	F50
AV9289	13,58	F50	14,08	F50	14,08	pSencillo	F50
AV9290	14,92	F50	14,42	F50	14,42	pSencillo	F50
AV9031	17,42	F50	17,92	F50	17,92	pSencillo	A318
AV9032	18,83	F50	18,33	F50	18,33	pSencillo	F50
AV9287	19,08	F50	19,58	F50	19,58	pSencillo	F50
AV9288	20,50	F50	20,00	F50	20,00	pSencillo	A318
AV8503	6,08	A318	6,58	A320	6,58	pSencillo	A320
AV8504	6,00	A318	5,50	A320	6,00	pSencillo	A319
AV9905	7,83	F50	8,33	A320	8,33	pSencillo	A320
AV8502	7,58	A318	7,08	A320	7,08	pSencillo	A319
AV8511	9,42	A318	9,92	A319	9,92	pSencillo	A319
AV9906	9,00	F50	8,50	A320	8,50	pSencillo	A320
AV8515	12,25	A318	12,75	A320	12,75	pSencillo	A320
AV8512	10,92	A318	10,42	A320	10,42	pSencillo	A320
AV8513	14,75	A318	15,25	A320	15,25	pSencillo	A318
AV8516	13,67	A318	13,17	A319	13,17	pSencillo	A319
AV8517	15,75	A318	16,25	A319	16,25	pSencillo	A319
AV8514	16,17	A318	15,67	A320	15,67	pSencillo	A320
AV9911	18,08	F50	18,58	A318	18,58	pSencillo	A320
AV8518	17,25	A318	16,75	A320	16,75	pSencillo	A318
AV9903	19,33	F50	19,83	A320	19,83	pSencillo	A319
AV9918	18,67	A320	18,17	A319	18,17	pSencillo	A319
AV9917	20,83	A320	21,33	A318	21,33	pSencillo	A319
AV9912	19,58	F50	19,08	A318	19,08	pSencillo	A319

AV8505	22,00	A318	22,50	A320	22,50	pSencillo	A319
AV9904	20,75	F50	20,25	A318	20,25	pSencillo	A320
AV9241	8,17	F50	8,67	F50	8,67	pSencillo	F50
AV9242	9,92	F50	9,42	F50	9,42	pSencillo	F50
AV9243	15,92	F50	16,42	F50	16,42	pSencillo	F50
AV9244	17,67	F50	17,17	F50	17,17	pSencillo	F50
AV8583	8,25	A318	8,75	A318	8,75	pSencillo	A318
AV8584	10,08	A318	9,58	A318	9,58	pSencillo	A318
AV9401	11,92	F50	12,42	A318	12,42	pSencillo	A318
AV9402	13,75	F50	13,25	A318	13,25	pSencillo	A318
AV8597	14,88	A318	15,38	A318	15,38	pSencillo	A318
AV8598	16,75	A318	16,25	A318	16,75	pSencillo	A318
AV58	7,50	A320	8,00	B767	8,00	pDoble	B767
AV59	10,00	A320	9,50	B767	10,00	pDoble	B767
AV60	14,42	B767	14,92	A330	14,92	pDoble	B767
AV61	17,92	B767	17,42	A330	17,42	pDoble	B767
AV82	20,75	A320	21,25	A320	21,25	pSencillo	A320
AV83	1,32	A320	1,82	A320	1,82	pSencillo	A320
AV8590	10,92	A318	11,42	A318	11,42	pSencillo	A318
AV8591	13,00	A318	12,50	A318	12,50	pSencillo	A318
AV9412	10,92	A318	11,42	A318	11,42	pSencillo	A318
AV9413	13,00	A318	12,50	A318	12,50	pSencillo	A318
AV94	11,92	A319	12,42	A319	12,42	pSencillo	A319
AV95	14,92	A319	14,92	A319	14,42	pSencillo	A319
AV97	21,83	A319	22,33	A330	22,33	pDoble	B767
AV98	8,42	A319	7,92	A330	7,92	pDoble	B767
AV250	14,00	A319	14,50	A320	14,00	pDoble	A330
AV251	18,50	A319	18,50	A320	18,00	pDoble	A330
AV8484	7,25	A318	7,75	A320	7,75	pSencillo	A320
AV8485	9,25	A318	8,75	A320	8,75	pSencillo	A320
AV9774	11,25	A320	11,75	A319	11,75	pSencillo	A320
AV9775	13,25	A320	12,75	A320	12,75	pSencillo	A320
AV9766	15,27	A320	15,77	A320	15,77	pSencillo	A318
AV9767	17,35	A320	17,35	A320	16,85	pSencillo	A320
AV9768	19,42	A320	19,92	A320	19,92	pSencillo	A320
AV9769	21,42	A320	20,92	A319	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV6371	8,75	A320	9,25	A320	9,25	pSencillo	A319
AV6378	6,33	A320	5,83	A320	5,83	pSencillo	A319
AV6373	14,17	A320	14,67	A319	14,67	pSencillo	A319

AV6372	11,38	A320	10,88	A320	11,38	pSencillo	A319
AV6375	18,75	A320	19,25	A320	19,25	pSencillo	A319
AV6374	16,33	A320	15,83	A319	15,83	pSencillo	A319
AV6377	22,00	A320	22,50	A319	22,50	pSencillo	A319
AV6376	18,83	A320	18,33	A319	18,33	pSencillo	A319
AV8596	8,83	A318	9,33	A319	9,33	pSencillo	A319
AV8595	10,83	A318	10,33	A319	10,33	pSencillo	A319
AV9410	12,58	A318	13,08	A320	13,08	pSencillo	A319
AV9411	14,42	A318	13,92	A319	13,92	pSencillo	A319
AV8592	17,75	A318	18,25	A319	18,25	pSencillo	A319
AV8593	19,67	A318	19,17	A320	19,17	pSencillo	A319
AV14	17,92	A330	18,42	A330	18,42	pDoble	A330
AV15	13,50	A330	13,00	A330	13,00	pDoble	A330
AV9433	6,00	F50	5,50	A320	6,50	pSencillo	A320
AV9434	6,08	F50	5,58	A320	5,58	pSencillo	A320
AV8525	6,92	A318	6,42	A320	7,42	pSencillo	A319
AV8524	8,17	A318	8,17	A318	8,17	pSencillo	A320
AV9435	12,92	F50	12,92	A320	12,42	pSencillo	A320
AV9033	14,35	A319	13,85	A320	14,35	pSencillo	A320
AV9431	18,92	F50	18,92	A320	18,42	pSencillo	A320
AV9430	15,17	F50	15,17	A319	15,17	pSencillo	A319
AV9034	21,92	A319	22,42	A320	22,42	pSencillo	A320
AV9432	20,33	F50	20,33	A320	20,33	pSencillo	A320
AV38	10,25	A320	10,75	A320	9,75	pSencillo	A320
AV39	17,35	A320	16,85	A320	16,85	pSencillo	A320
AV9403	10,50	F50	10,50	F50	11,00	pSencillo	A318
AV9404	11,70	F50	11,70	F50	11,20	pSencillo	A318
AV9405	16,50	F50	16,50	F50	17,00	pSencillo	A318
AV9406	17,75	F50	17,75	F50	17,75	pSencillo	A318
AV9419	7,63	F50	7,63	F50	8,13	pSencillo	A318
AV9420	9,17	F50	9,17	F50	8,67	pSencillo	A318
AV9761	9,00	F50	8,50	A318	9,00	pSencillo	A318
AV9760	7,33	F50	7,83	A318	6,83	pSencillo	A318
AV9511	16,08	F50	16,08	A318	16,58	pSencillo	F50
AV9510	14,50	F50	Vuelo no elegido	Vuelo no elegido	14,50	pSencillo	F50
AV9771	20,08	F50	20,08	A318	20,58	pSencillo	F50
AV9770	18,50	F50	18,50	A318	19,00	pSencillo	F50
AV34	9,17	A320	9,17	A319	8,67	pSencillo	A319
AV35	14,17	A320	13,67	A319	14,67	pSencillo	A319
AV9035	16,83	A320	17,33	A318	16,83	pSencillo	A319

AV9036	7,00	A320	6,50	A318	6,50	pSencillo	A319
AV43	9,00	A319	8,50	A319	8,50	pSencillo	A319
AV42	23,92	A319	23,92	A319	23,92	pSencillo	A319
AV17	13,17	A330	12,67	A330	12,67	pDoble	A330
AV16	18,33	A330	18,83	A330	18,83	pDoble	A330
AV30	10,00	A320	10,50	A320	10,00	pSencillo	A320
AV31	15,75	A320	15,25	A320	15,25	pSencillo	A320
AV9360	10,67	F50	10,67	F50	Vuelo no elegido	Vuelo no elegido	Vuelo no elegido
AV9361	12,50	F50	13,00	F50	12,50	pSencillo	A318

D. Anexo: Código en R del algoritmo de clusterización de datos

```
#Lectura de datos
datos=read.table(file.choose(),sep=",")

#Conteo del número de observaciones
n=length(datos[,1])

#Crear octava columna con ceros (será la columna con el ClusterID) y
#novena columna con unos (para monitorear el cambio de cluster)
datos[,8]=c(rep(0,n))
datos[,9]=c(rep(1,n))

#Nombrar las columnas de los datos
colnames(datos)=c("CASM","Sillas","Costo operativo","Alcance",
                  "P_desp", "P_ater","Nombre","ClusterID","CChange")

#Guardar los valores maximo y mínimo de cada criterio técnico y normalizar los datos
minimo=c()
maximo=c()
for(i in 1:6){
  minimo[i]=min(datos[,i])
  maximo[i]=max(datos[,i])
  datos[,i]=(datos[,i]-min(datos[,i]))/(max(datos[,i])-min(datos[,i]))
}

#Número de clusters
k=2

#Creación de centroides (aleatorios)
centroides=list()
for(i in 1:k){
  centroides[[i]]=runif(n=6,min=0,max=1)
}

#Función para calcular distancia con métrica 2
dist=function(a,b){
  distancia=0
  if(length(a)==length(b)){
    dim=length(a)
    for(i in 1:dim){
      distancia=distancia+(a[i]-b[i])^2
    }
    distancia=sqrt(distancia)
  }
```

```

    return(distancia)
  }
}

#Inicio del proceso iterativo
niteraciones=0

#Clusterizacion
while(sum(datos["CChange"])!=0){
#Asignacion de cluster a cada punto
for (i in 1:n) {
  distancia = c()
  for (j in 1:k) {
    distancia[j] = dist(datos[i, 1:6], centroides[[j]])
  }
  cluster = which.min(distancia)
  if (cluster != datos[i, 8]) {
    datos[i, 8] = cluster
    datos[i, 9] = 1
  } else {
    datos[i, 9] = 0
  }
}
#Recálculo de los centroides
for(j in 1:k){
  suma=c(0,0,0,0,0,0)
  nobs=0
  for(i in 1:n){
    if(datos[i,"ClusterID"]==j){
      suma=suma+datos[i,1:6]
      nobs=nobs+1
    }
  }
  centroides[[j]]=suma/nobs
}
niteraciones=niteraciones+1
}

#Mostrar los centroides definitivos sin escala
centroides

#reescalar los centroides
for(i in 1:k){
  centroides[[i]]=centroides[[i]]*(maximo-minimo)+minimo
}

#Mostrar los centroides definitivos en su escala original
centroides

```